

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 10:52:04 ; Search time 1705 Seconds
(without alignments)
4702.165 Million cell updates/sec

Title: US-09-369-992c-l_COPY_1147_1740
Perfect score: 594
Sequence: 1 gtagcgtttaatagcgaa.....ttacatctttcagcggtc 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175.4	29.5	628	9	AL671944
C 2	171.4	28.9	571	9	BE230849
C 3	171.4	28.9	598	9	AI151682
C 4	170.8	28.8	636	12	AZ525208
C 5	169	28.5	666	12	BH504408
C 6	169	28.5	671	12	BH466711
C 7	169	28.5	674	12	BH425423
C 8	169	28.5	684	12	BH562802
C 9	169	28.5	690	12	BH504251
C 10	169	28.5	696	12	BH472129
C 11	169	28.5	702	12	BH446986
C 12	169	28.5	704	12	BH466757
C 13	169	28.5	709	12	BH571671
C 14	169	28.5	721	12	BH537063
C 15	169	28.5	725	12	BH448906
C 16	169	28.5	725	12	BH605990
C 17	169	28.5	726	12	BH471550

C 18	169	28.5	729	12	BH569297
C 19	169	28.5	735	12	BH539628
C 20	169	28.5	737	12	BH430587
C 21	169	28.5	737	12	BH479773
C 22	169	28.5	737	12	BH581183
C 23	169	28.5	739	12	BH480050
C 24	169	28.5	741	12	BH568000
C 25	169	28.5	744	12	BH436695
C 26	169	28.5	750	12	BH545017
C 27	169	28.5	752	12	BH509024
C 28	169	28.5	757	12	BH491792
C 29	169	28.5	758	12	BH507404
C 30	169	28.5	758	12	BH572558
C 31	169	28.5	758	12	BH584910
C 32	169	28.5	759	12	BH608306
C 33	169	28.5	760	12	BH452416
C 34	169	28.5	762	12	BH475346
C 35	169	28.5	762	12	BH505034
C 36	169	28.5	765	12	BH469296
C 37	169	28.5	765	12	BH474848
C 38	169	28.5	769	12	BH608446
C 39	169	28.5	771	12	BH598135
C 40	169	28.5	773	12	BH569181
C 41	169	28.5	776	12	BH499158
C 42	169	28.5	776	12	BH570901
C 43	169	28.5	781	12	BH512342
C 44	169	28.5	783	12	BH543171
C 45	169	28.5	784	12	BH587376

ALIGNMENTS

RESULT 1
AL671944
LOCUS AL671944 Scherffellia dubia LambdaZAPII linear EST 08-FEB-2002
DEFINITION p57, mRNA sequence.
ACCESSION AL671944
VERSION AL671944.1 GI:18641526
KEYWORDS EST.
SOURCE Scherffellia dubia.
ORGANISM Scherffellia dubia
Eukaryota; Viridiplantae; Chlorodendrales; Chlorodendraceae; Scherffellia.
REFERENCE 1 (bases 1 to 628)
AUTHORS Becker, B., Feja, N. and Melkonian, M.
TITLE Analysis of expressed sequence tags (ESTs) from the scaly green flagellate Scherffellia dubia Pascher emend. Melkonian et Preisig
JOURNAL Protist 152 (2), 139-147 (2001)
MEDLINE 21428165
COMMENT Contact: Becker B
Botanisches Institut
Universitaet zu Koeln
Gyrhofstr. 15, 50931 Koeln, Germany.
Location/Qualifiers
1. 628
/organism="Scherffellia dubia"
/strain="SAG 40.89"
/db_xref="taxon:3190"
/clone="p57"
/clone_lib="Scherffellia dubia LambdaZAPII"
BASE COUNT 146 a 185 c 131 g 165 t 1 others
ORIGIN

Query Match 29.5%; Score 175.4; DB 9; Length 628;
Best Local Similarity 60.5%; Pred. No. 7.2e-17;
Matches 370; Conservative 0; Mismatches 217; Indels 25; Gaps 4;
QY 4 TCCTTTAATAGCGGACAGACTTACCTTTAAACATACACTACTGCTTAGGATGCGATAA 63
||||||| ||||||||| ||||| ||||||| || ||||| ||
DB 7 TCCTTTTCATGGCGAACAG--CAACCTTGGGACGTACTACGCCCCCGAGTTGCGAGAA 64

TITLE Genes expressed in adult female stage of Onchocerca volvulus
JOURNAL Unpublished (1998)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genomesmith.edu
Seq primer: pbluescript SK.
FEATURES Location/Qualifiers
source
1..598
/organism="Onchocerca volvulus"
/db_xref="taxon:6282"
/clone="SWOAFCAP07B11"
/clone_lib="Onchocerca volvulus adult female cDNA
(SAW98MLW-OVAF)"
/sex="female"
/dev_stage="adult"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda Uni-ZAP XR; Site.1: Eco RI; Site.2:
Xho I; Filarial nematode parasite of humans. Two adult
female worms of Onchocerca volvulus were isolated from
consenting patients and quick frozen. Adult female mRNA
was converted to double-stranded cDNA using reverse
transcriptase and oligo(dT) followed by RNase H and DNA
pol I. The library has 7 x 10E5 independent recombinants
and the average insert size is ~1100bp. The library was
constructed by Michelle Lizotte-Waniewski with worms
provided by Dr. Sara Lustigman. The library is available
from Dr. Steven A. Williams, email: genomesmith.edu."

BASE COUNT 154 a 124 c 166 g 151 t
ORIGIN
Query Match 28.9%; Score 171.4; DB 9; Length 598;
Best Local Similarity 60.9%; Pred. No. 2.8e-16;
Matches 357; Conservative 0; Mismatches 204; Indels 25; Gaps 4;
QY 1 GTATCGCTTTAATAGCGAACACATTACCTTTAAACATCTACTACTCTTAGGATGCGA 60
DB 588 GTATCACTTTATCGGANGANCAGCAACCTT-GGACCTCTTTCAGCCCCAGAGTGTGA 530
QY 61 TAAGCCGACATCGAGTGCACCAACCTTT-TCGTCAATATGACTCTCGGAAAGATTAGC 119
DB 529 TGAGTCGACATCGAGTGCACCAAGGGTGTGTCGTGATATGAACCTTTGAACACCATCAGC 470
QY 120 CTGTTATCCCTAGAGTAACCTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATC 179
DB 469 CTGTTATCCCGCGGTACCTTTTATCCGTTGAGCGATGACCCCT-TCATACAGAAATCACC 411
QY 180 GGATCATTAAGACGACATTAATCTCTGTTTAAATTTTACAGTTAATTATATATATA 239
DB 410 GGATCACTATGACGACTTCGTCTCTGCTTGTGCGCTCGACGTCGAGCGAAGCT 351
QY 240 TTTATCTTTTATAATAA-----TATAACATTGTACACTCC 277
DB 350 TGTGCATTATCTATTAAAGCTGATTTCCGACACGCTCTAGCTTACCTTCGACGCTCC 291
QY 278 GTTTTATATAGAGAGACCGCCGAGCAACTATCTTATAATATTGTTAAAAATTT 337
DB 290 GTTACCTTTTAGGCGCGACCGCCCGAGTCAAACTACCCACCATACAATGTCTAGTCC 231
QY 338 TGTATAAATAATTTTATAAGAAATTTATATATATAATAAATGGTATTTCATTAAACATTAC 397
DB 230 AGATATGAACACATAGTTAGATATACAAAGTGAAGGGTGGTATCTCAAGGTGCACTCC 171
QY 398 ATTATTTCCAAAAAATAATATTACTACTTCCCATTTTATCTATGTTATATATATATAT 457
DB 170 ATTACAGCTAGCGCCATAACTTCAAAGCTCCCACTATCTCTGCACATTTACATTTTAT 111
QY 458 TTCATATCTATTATAGTAAGAGCTTCATAGGGTCTTCTGTGCCCAATATATAGAAATCTG 517

DB 110 AGCAATATAAAGCTATAGTAAAGTGCACGGGTCTCTTCGTCTAACCGGGGTACCCCG 51
QY 518 CATCTTCACAGATAAATTTTATTCATTAAGATTTTTTTTAAGACAG 563
DB 50 CATCTGCACGGGAATTCAAATTCGCTGAATGTGATGGAGACAG 5
RESULT 4
AZ525208/c 636 bp DNA linear GSS 07-MAY-2001
LOCUS 241PBD09 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.
DEFINITION
ACCESSION AZ525208
VERSION AZ525208.1 GI:13965828
KEYWORDS
SOURCE Plasmodium berghei.
ORGANISM Plasmodium berghei
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
AUTHORS 1 (bases 1 to 636)
CARLTON,J.M.-R. and DAME,J.B.
TITLE The plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL Parasitol. Today 16 (10), 409 (2000)
COMMENT Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.
FEATURES Location/Qualifiers
source
1..636
/organism="Plasmodium berghei"
/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
/db_xref="taxon:5821"
/clone_lib="Pb MBN #21"
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
/note="Vector: pBluescript SK(+), vector DNA, phagemid
excised from lambda ZAP; Site.1: EcoRV; Site.2: EcoRV;
Genomic DNA was prepared from asynchronous blood stage
forms of the cloned ANKA isolate of P. berghei grown in
laboratory Swiss white mice. The DNA was purified from
contaminating host DNA by Hoechst dye 33258-CsCl
ultracentrifugation and precipitated. Purified DNA was
digested with mung bean nuclease in the presence of 36-38%
formamide at 50 C, as described (Vernick, K.D., Imberski,
R.B., and McCutchan, T.F. 1988. Nucleic Acids Research
16:6883-6896). The ends of the digestion fragments were
polished using T4 DNA polymerase, and the fragments size
selected in the range 500-2000 bp. These were ligated into
the EcoRV-cleaved and dephosphorylated pBluescript SK(+)
vector. Recombinant plasmids were used to transform E.
coli XL10-Gold host cells."

BASE COUNT 261 a 63 c 88 g 223 t 1 others
ORIGIN
Query Match 28.8%; Score 170.8; DB 12; Length 636;
Best Local Similarity 93.1%; Pred. No. 3.4e-16;
Matches 201; Conservative 0; Mismatches 12; Indels 3; Gaps 2;
QY 382 TTTTCATTAAACAATTATTTCC--AAAAAATAATATTACTACTTCCCA-TTTATTTC 438
DB 634 TTTTCATTAAACAATTATTTTCGAAAAAATAATATGACGCTTCCCATTTATTC 575
QY 439 TAGTATATATATATATTTTCAATATCTATTAATAGTAAGCTTCATAGGTCCTTCG 498
DB 574 TAGTATATAGATATATTTTCAATATCTATTAATAAGTAAGCTTCATAGGTCCTTCG 515
QY 499 TCCATATATAAGAAATCTGCATCTTCACAGATAATTTTATTTCATTAAAGATTTTTTAA 558
DB 514 TCATAATAAGAANAATCTGCATCTTCACAGATAATTTTATTTCATTAAAGATTTTTTAA 455

```
QY 559 GACAGCATTTAAGTCGTTACATCTTTTCATCGAGTGC 594
|||||
Db 454 GACAGCATTTAAGTCGTTACATCTTTTCATCGAGTGC 419

RESULT 5
BH504408/c
LOCUS
DEFINITION BH504408 666 bp DNA linear GSS 13-DEC-2001
BOHF234TR BOHF Brassica oleracea genomic clone BOHF234, DNA
sequence.
ACCESSION BH504408
VERSION BH504408.1 GI:117712505
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 666)
TOWN,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOHF234TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
Source
1..666
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHF234"
/clone_lib="BOHF"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 152 a 165 c 213 g 136 t
ORIGIN
Query Match 28.5%; Score 169; DB 12; Length 666;
Best Local Similarity 58.6%; Pred. No. 6.1e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACTACTGCGCTTAGGATGCGA 60
Db 642 GTACCGCTTTAATAGCGGAACAGCCCAACCCCTTGAACATACTACTACAGCCCGAGTGGCGA 583

QY 61 TAAGCCGACATCGAGTGCAGAACCCCTTTCGTCGAATATGACCTCTCGGAAAGATTAGCC 120
|||||
Db 582 AGAGCCGACATCGAGTGCAGAACCCCTTTCGTCGATGTGAGCTCTTGGGGAAGATCAGCC 523

QY 121 TGTATCCCTAGACTAACTTTTATCCCTTAAGCGATAATTTTATTAATAATTAATATCG 180
|||||
Db 522 TGTATCCCTAGACTAACTTTTATCCCTTGAGCGAGCGCCCTTCCACTCGGACCGTCG 464

QY 181 GATCATTAAAGACCGACATTAATCTCTGTTTAAATTTGTAATTTTACAGATTAATATATAT 240
|||||
Db 463 GATCATTAAAGCGGACCTTTCGTCGCCCTGCTCGAGGGTGGGTCTTGCAGCTCAAGCTCCCTT 404

QY 241 TTATCTTTATATATA-----ATATAACATTTGTACACCTCCG 278
|||||
Db 403 CTGCTTTTGCACCTCGAGGGCCAATCTCCGTCGCGCCCGAGGAACCTTTGACGCGCTCCG 344

QY 279 TTTTATATAGAGGAGACCGCCCACTCAAACTATCTTTATAAATATTGT---TAAAAAT 335
|||||
Db 343 TTACCTTTTGGAGGCTACGCCCCATAGAACTGTCTACCTGAGACTGTCCCTTGGCCC 284

QY 336 TTTGTTATAAAAAATTTTATAAGAAATTTATATATATAATAAATGGTATTTTCATTAACAATT 395
|||||
```

```
Db 283 GTAGGTCCTGACACAAGGTTAGAATTTAGCTCTCTCCAGAGTGGTATCTCACTGATGGCT 224
|||||
QY 396 ACATTATTTCCAAAAAATAATATTTACTACTTCCCATTTATTTCTATGTTATATATATA 455
|||||
Db 223 CGGGCCCCCGGAGGAGGCCCTTCTTCGCCCTTCCACCTAAGCTGCGCAGGAAAGCCCA 164
|||||
QY 456 TTTTCAATATCTAATTAATAGCTTTCATAGGCTTTCTAGGCTTTCTTCCTCAATAATAAGAAATC 515
|||||
Db 163 AAGCCAATCCAGGACAGTGAAGCTTTCATAGGCTTTCTGTCAGGTCGAGGTAGTC 104
|||||
QY 516 TGCATCTTCACAGATAATTTTATTTTATTCAATTAAGATTTTTTTTAAGACGACATTTAAGTCGT 575
|||||
Db 103 CGCATCTTCACAGACATGTCTATTTTACCGAGCGCTCTCTCCGAGACAGTGCAGATCGT 44
|||||
QY 576 TACATCTTTCATCGAGTGC 594
|||||
Db 43 TAGCGCTTTTCGTGCGGGTC 25

RESULT 6
BH466711/c
LOCUS
DEFINITION BH466711 671 bp DNA linear GSS 13-DEC-2001
BOGRN92TR BOGR Brassica oleracea genomic clone BOGRN92, DNA
sequence.
ACCESSION BH466711
VERSION BH466711.1 GI:17667845
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 671)
TOWN,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOGRN92TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
Source
1..671
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGRN92"
/clone_lib="BOGR"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 152 a 165 c 214 g 140 t
ORIGIN
Query Match 28.5%; Score 169; DB 12; Length 671;
Best Local Similarity 58.6%; Pred. No. 6.1e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACTACTGCGCTTAGGATGCGA 60
|||||
Db 655 GTACCGCTTTAATAGCGGAACAGCCCAACCCCTTGAACATACTACTACAGCCCGAGTGGCGA 596

QY 61 TAAGCCGACATCGAGTGCAGAACCCCTTTCGTCGAATATGAGCTCTCGGAAAGATTAGCC 120
|||||
Db 595 AGAGCCGACATCGAGTGCAGAACCCCTTTCGTCGATGTGAGCTCTTGGGGAAGATCAGCC 536

QY 121 TGTATCCCTAGACTAACTTTTATCCGTTAAGCGATAATTTTATTAATAATTAATATTCG 180
|||||
```

```

Db 535 TGTTATCCCTAGAGTAACCTTTTATCCGTTGAGCGACGGCCCTTCCACT -CGGCACCGCTCG 477
QY 181 GATCATTAAAGACGACATTAATCTCTGTTTAAATTTGTAAATTTTACAGTTAAATTATATAT 240
Db 476 GATCATAAGGCGAGCTTTCGTCCTCGTCTGACGGGTGGTCTTGAGTCAAGCTCCCTT 417
QY 241 TTATCTTTTATATATA-----ATATAACATTTGTACACCTCCG 278
Db 416 CTGCTTTTGCACTCGAGGGCCAATCTCCGTCGCGCCCGAGGAAACCTTTGCACGCCCTCG 357
QY 279 TTTTATATAGGAGAGACGCGCCAGTCAAACTATCTTATAATATTGT -TAAAAAT 335
Db 356 TTACCTTTTGGAGGCGCTACGCGCCATAGAAATGTCTACCTGAGACTGCCCTTGGCCC 297
QY 336 TTTGTTATATAAATTTTATAAGAAATTTATATATATAATAAATGGTATTTTCATTAAACAAT 395
Db 296 GTAGGTCCTGACACAGGTTAGAAATCTAGCTCTCCAGAGTGGTATCTCACTGATGGCT 237
QY 396 ACATTAATTTCCAAAAAATAATTAATCTACTTCCCAATTTATCTATGTTATATATATATA 455
Db 236 CGGGCCCCCGGAGAGGCGCTTCTCGCTTCCACCTAAGCTGCGCAGGAAAGCCCA 177
QY 456 TTTTCAATATCTAATATAGTAAGCTTCATAGGGTCTTCTCTCTCTAATAATAAGAAATC 515
Db 176 AAGCCAAATCCAGGGAACAGTGAAGCTTCATAGGGTCTTCTGTCAGGTGCAAGTAGTC 117
QY 516 TGCATCTTCACAGATAAATTTTATTTTATTAAGATTTTAAAGACAGCATTTTAAGTCGT 575
Db 116 CGCATCTTCACAGACATGTCTATTTCACCGAGCTCTCTCCGAGAGTGCAGCCAGATCGT 57
QY 576 TACATCTTCATGCAAGTC 594
Db 56 TACGCTTTTCGTGCGGGTC 38

RESULT 7
BH425423
LOCUS BH425423R BOHE Brassica oleracea genomic clone BOHEL23, DNA
DEFINITION sequence.
ACCESSION BH425423
VERSION BH425423.1 GI:17611151
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
COMMENT Whole genome shotgun sequencing of Brassica oleracea
Other GSSs: BOHEL23TF
Contact: Chris Town
TIGR
199 a 214 c 167 g 154 t
BASE COUNT 139 a 214 c 167 g 154 t
ORIGIN

```

```

Query Match 28.5%; Score 169; DB 12; Length 674;
Best Local Similarity 58.6%; Pred. No. 6e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGGCGAAGACACTTTACCTTTAAAAACATACTACTGCTTACGATGCGA 60
Db 44 GTACCGCTTTAATGGGGAACAGCCCAACCTTTGGAACATACATACAGCCCGAGTGGCGA 103
QY 61 TAGCCGACATCGAGGTGCGCAACCTTTTCGTCATATGAGACTCTCGGAAAGAAATTAAGC 120
Db 104 AGACCGGACATCGAGGTGCGCAACCTTTCCGTCGATGTGAGCTCTTGGGGAAGATCAGCC 163
QY 121 TGTATATCCCTAGAGTAACTTTTATCCGTTAAGCGATAAATTTTATTAATAATTAATTA 180
Db 164 TGTATATCCCTAGAGTAACTTTTATCCGTTGAGCGACGGCCCTTCCACT -CGGCACCGCTCG 222
QY 181 GATCATTAAAGACGACATTAATCTCTGTTTAAATTTTGTAAATTTTACAGTTAAATTATAT 240
Db 223 GATCATAAGGCGAGCTTTTCGTCCTGCTGACGGGTGGTCTTGCAGTCAAGCTCCCTT 282
QY 241 TTATCTTTTATATAA-----AATATAACATTTGTACACCTCCG 278
Db 283 CTGCTTTTGCACTCGAGGGCCAATCTCCGTCGCGCCCGAGGAAACCTTTGCACGCCCTCG 342
QY 279 TTTTATATAGGAGAGACGCGCCAGTCAAACTATCTTATAAATATTGT -TAAAAAT 335
Db 343 TTACCTTTTGGGAGGCGCTACGCGCCATAGAACTGTCTACTGAGACTGTCCTTGGGCC 402
QY 336 TTTGTTATAAATTTTATAAGAAATTTATATATATATAAATTTTATTAATTTTCAATTAACAAT 395
Db 403 GTAGGTCCTGACACAGGTTAGAAATCTAGCTCTTCCAGAGTGGTATCTCACTGATGGCT 462
QY 396 ACATTAATTTCCAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 455
Db 463 CGGGCCCCCGGAGAGGAGGCTTCTTCGCTTCCACTAAGCTGCGCAGGAAAGCCCA 522
QY 456 TTTTCAATATCTAATTAAGTAAAGCTTCATAGGGTCTTCTGTCCTTAATAATAAGAAATC 515
Db 523 AAGCCAAATCCAGGAGACAGTGAAGCTTCATAGGGTCTTCTGTCAGAGTGCAGGTAGTC 582
QY 516 TGCATCTTCACAGATAAATTTTATTAAGATTTTATTAAGATTTTATTAAGACAGCATTTAAGTCGT 575
Db 583 CGCATCTTCACAGACATGTCTATTTCACCGAGCTCTCTCCGAGAGACTGCCAGATCGT 642
QY 576 TACATCTTCATGCAAGTC 594
Db 643 TACGCTTTTCGTGCGGGTC 661

RESULT 8
BH562802
LOCUS BH562802R BOGM Brassica oleracea genomic clone BOGMW42, DNA
DEFINITION sequence.
ACCESSION BH562802
VERSION BH562802.1 GI:17814642
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
COMMENT Whole genome shotgun sequencing of Brassica oleracea
Other GSSs: BOGMW42TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

```



```
QY 576 TACATCTTTTCATCAGGTC 594
    ||| |||| ||| ||||
Db 64 TACGCGCTTTCGTGGGTC 46

RESULT 10
BH472129
LOCUS BH472129 696 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGQ90TF BOGQ Brassica oleracea genomic clone BOGQ90, DNA
sequence.
ACCESSION BH472129
VERSION BH472129.1 GI:17680240
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 696)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other-GSSs: BOGQ90TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..696
/organism="Brassica oleracea"
/strain="TO1000H3"
/db_xref="taxon:3712"
/clone="BOGQ90"
/clone_lib="BOGQ"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 144 a 222 c 171 g 159 t
ORIGIN

Query Match 28.5%; Score 169; DB 12; Length 696;
Best Local Similarity 58.6%; Pred. No. 6e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCTTAAACATACACTACTGCCTTAGGATGCGA 60
Db 41 GTACCGCTTTAATGGCGAACAGCCCAACCCCTTGGAAACATACACTACAGCCCGAGTGGCGA 100
QY 61 TAAGCGGACATCGAGTGGCCAAACCTTTTCGTCATATAGACTCTCGGAAAGATTAGCC 120
Db 101 AGAGCGGACATCGAGTGGCCAAACCTTCCCGTCGATGAGCTCTGGGGAAGATCAGCC 160
QY 121 TGTATPCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
Db 161 TGTATCCCTAGAGTAACCTTTTATCGTTGAGCGAGCGCCCTTCCACT-CGGCACCGTCG 219
QY 181 GATCAATAAGCGACGACATTAATCTCTGTTTAATTTGTAATAATTACAGTTAATATATAT 240
Db 220 GATCAATAAGCGGCGACTTTTCGTCCTCGTCGACGGGTGGTCTTGCACTCAAGCTCCCTT 279
QY 241 TTATCTTTTATATAA-
Db 280 CTGCGCTTCGCACTAGAGGCGCAATCTCCGTCGCGCCCGAGGAACCTTTGCACGCGCTCG 339
QY 279 TTTTATATAGGAGAGACGCGCCAGTCAAACTATCTTTATATAATATTGCT---TAAAAAT 335
Db 340 TTACCTTTTGGAGGCGCTTACGCCCATAGAACTGTCTACCTGAGACTGTCCCTTGGCGCC 399
```

```
QY 336 TTTGTTATAAAAAATTTTATAAGAATTTATATATATAATAAAATGTTATTTCAATTAACAATT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 GTAGGTCCTGCACACAAGTTAGATTCTAGCTCTTCAGAGTGGTATCTCACTGATGGCT 459
QY 396 ACATTATTTCCAAAAAATAATATTAATACTACTTCCCATTTTATTTCTATATATATATA 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 CGGGCCCCCGGAAGGAGGCGCTTCTTCGCGCTTCCCACTAAGCTGCGCAGAAAGGCCA 519
QY 456 TTTTCAATATCTATTATACTAAAGCTTCATAGGCTCTTCTCTGCTCCCTCAATATAGAATC 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 AAGCAATCCAGGAACAGTGAAGCTTCATAGGCTCTTCTGCTCCAGGTGCGAGGTAGTC 579
QY 516 TGCATCTTCACAGATAATTTTATTTCAATTAAGATTTTTTTTAAGACAGCATTTTAAGTCGT 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 CGCATCTTCACAGACATGTCTATTTTACCGAGGCTCTCTCCGAGACAGTCCCGAGATCGT 639
QY 576 TACATCTTTCATCAGGTC 594
    ||| |||| ||| ||||
Db 640 TACGCGCTTTCGTGGGTC 658

RESULT 11
BH446986
LOCUS BH446986 702 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGDC64TF BOGD Brassica oleracea genomic clone BOGDC64, DNA
sequence.
ACCESSION BH446986
VERSION BH446986.1 GI:17632700
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 702)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other-GSSs: BOGDC64TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..702
/organism="Brassica oleracea"
/strain="TO1000H3"
/db_xref="taxon:3712"
/clone="BOGDC64"
/clone_lib="BOGD"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 139 a 222 c 179 g 162 t
ORIGIN

Query Match 28.5%; Score 169; DB 12; Length 702;
Best Local Similarity 58.6%; Pred. No. 6e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCTTAAACATACACTACTGCCTTAGGATGCGA 60
Db 1 GTACCGCTTTAATGGCGAACAGCCCAACCCCTTGGAAACATACACTACAGCCCGAGTGGCGA 60
QY 61 TAAGCGGACATCGAGTGGCCAAACCTTTTCGTCATATAGACTCTCGGAAAGATTAGCC 120
Db 61 AGAGCGGACATCGAGTGGCCAAACCTTCCCGTCGATGAGCTCTTGCGGAAGATCAGCC 120
QY 121 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
```

```

121 TGTATCCCTAGAGTAACATTTTATCCGTTGAGCGACGGCCCTCCACT-CGGCACCGTCG 179
181 GATCATTAAGACCCACACATTAATCTCTCTTTAAATTTGTAATTTTACAGTTAATATATAT 240
180 GATCATAAGGCCGACCTTCGTCCTCTGCTGACGGGTGGTCTTGACAGTCAAGTCCCTT 239
241 TTATCTTTATATAATA-----AATATAACATTTGTACACCTCCG 278
240 CTGCTTTTGCACTCGAGGGCCAACTCTCCGTCGGCCCGAGGAACCTTTGCACGCTCCG 299
279 TTTTATATAGGAGGACCGCCAGTCAAACTATCTTTATAAATATGT- -TAAAAAT 335
240 TTTTATATAGGAGGACCGCCAGTCAAACTATCTTTATAAATATGT- -TAAAAAT 335
300 TTACCTTTTGGGAGCCCTACGCCCCATAGAACTGTCTAAGTGTGAGTGTGCTTGGGCC 359
336 TTTCTTATAAATTTTATAGAAATTTATATATATATAAATGGTATTCATTAACAAT 395
360 GTAGGCTCTGACAAAGGTTAGAAATCTAGCTCTTCAGAGTGGTATCTCACTGATGGCT 419
396 ACATTAATTTCCAAAAAATAATATTAAGTATCTTCCCTTATCTATGTTATATATATA 455
420 CGGGCCCCCGGAGGAGGCTTCTTCGCTTCCACTTAAGCTGCGCAGGAAAGCCCA 479
456 TTTTCAATATCTATATAGTAAGCTTCATAGGCTTCTTCTGCTCAATATAAAGAAATC 515
480 AAGCAATCCCGAGGAAACAGTGAAGCTTCATAGGCTTCTTCTGCTCCAGGTGCAAGT 539
516 TGCATCTTCACAGATAATTTATTTCAATTAAGATTTTATTAAGACAGCATTTAAGTCG 575
540 GCGATCTTCACAGACATGTCTATTTTACCGAGGCTCTCTCCGAGACAGTGCCCGATCG 599
576 TACATCTTTTCATGCGAGTC 594
600 TACGCTTTTCGTGCGGTC 618

RESULT 12
BH466757/c
LOCUS
DEFINITION BH466757 704 bp DNA linear GSS 13-DEC-2001
BOHB08TR BOHB Brassica oleracea genomic clone BOHB08, DNA
sequence.
ACCESSION BH466757
VERSION BH466757.1 GI:17668010
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 704)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOHB08TF
Contact: Chris Town
TIGR
7912 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..704
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHB08"
/clone_lib="BOHB"
/note="vector: pHOS1; site:1; BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 161 a 175 c 220 g 148 t
```

```

ORIGIN
Query Match 28.5%; Score 169; DB 12; Length 704;
Best Local Similarity 58.6%; Pred. No. 6e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAAGACACTTACCTTTAAAACATACTACTGCTTGGATAGCGA 60
DB 660 GTACCGCTTTAATAGCGAAGACACTTACCTTTAAAACATACTACTGCTTGGATAGCGA 601
QY 61 TAGCCGACATCGAGGTGCGCAACCTTTTCGTCAAATATGGAATCTCGGAAAGATAGCC 120
DB 600 AGAGCCGACATCGAGGTGCGCAACCTTTTCGTCGATGTGAGCTCTTGGGGAAGATACCG 541
QY 121 TGTATCCCTTAGAGTAACTTTTATCCGTTAAGCGATAATTTTATTTAAATAATTATCG 180
DB 540 TGTATCCCTTAGAGTAACTTTTATCCGTTAAGCGATAATTTTATTTAAATAATTATCG 482
QY 181 GATCATTAAAGACGACATTAATCTCTTTTAAATTTGTAATTTTACAGTTAATATATAT 240
DB 481 GATCATTAAAGACGACATTTTCGTCGCTGCTGAGGGTGGTCTTGCAGTCAAGCTCCCT 422
QY 241 TTATCTTTATATAATA-----AATATAACATTTGTACACCTCCG 278
DB 421 CTGCTTTTGCACTCGAGGGCCAACTCTCCGTCGGCCCGAGGAACCTTTGCACGCTCCG 362
QY 279 TTTTATATAGGAGGACCGCCAGTCAAACTATCTTTATAAATATGT- -TAAAAAT 335
DB 361 TTACCTTTTGGGAGGCTACGCCCATAGAACTGTCTAAGTGTGCTCCAGGTGCAAGT 302
QY 336 TTTTATATAAATTTTATAGAAATTTATATATATAAATGGTATTTCAATTAACAAT 395
DB 301 GTAGGCTCTGACACAAAGTTAGAAATCTAGCTCTTCCAGAGTGTCTACTCATGAGCT 242
QY 396 ACATTAATTTCCAAAAAATAATATTAAGTATCTTCCCTTATCTATGTTATATATATA 455
DB 241 CGGGCCCCCGGAGGAGGCTTCTTCGCTTCCACTTAAGCTGCGCAGGAAAGCCCA 182
QY 456 TTTTCAATATCTATTAAGTAAAGCTTCATAGGCTTCTTCTGCTCAATATAAAGAAATC 515
DB 181 AAGCAATCCCGAGGACAGTGAAGCTTCATAGGCTTCTTCTGCTCCAGGTGCAAGT 122
QY 516 TGCATCTTCACAGATAATTTTATTTCAATTAAGATTTTATTAAGACAGCATTTAAGTCG 575
DB 121 CGCATCTTCACAGACATGTCTATTTTACCGAGGCTCTCTCCGAGACAGTGCCCGATCG 62
QY 576 TACATCTTTTCATGCGAGTC 594
DB 61 TACGCTTTTCGTGCGGTC 43

RESULT 13
BH571671
LOCUS
DEFINITION BH571671 709 bp DNA linear GSS 14-DEC-2001
BOHQZ22TF BOHQ Brassica oleracea genomic clone BOHQZ22, DNA
sequence.
ACCESSION BH571671
VERSION BH571671.1 GI:17823510
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 709)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOHQZ22TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
```


Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .709
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHQ222"
/clone_lib="BOHQ"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 151 a 226 c 171 g 161 t
ORIGIN

Query Match 28.5%; Score 169; DB 12; Length 709;
Best Local Similarity 58.6%; Pred. No. 6e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACACTACTGCTTACGCTTAGGATGCGA 60
DB 63 GTACCCCTTTAATGGCGAACAGCGCAACCCCTTGGAAACATACACTACAGCCCCAGGTCGCGA 122
QY 61 TAAGCGGACATCGAGTGGCGCAACACCTTTTCGCAATATGAGCTCTCTCGGAAAAGATTAGCC 120
DB 123 AGAGCGGACATCGAGTGGCGCAACACCTTTTCGCAATATGAGCTCTCTCGGAAAAGATTAGCC 182
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGGATAATTTTATTAATTAATTAATTCG 180
DB 183 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGGATAATTTTATTAATTAATTAATTCG 241
QY 181 GATCATTAAGACCGACATTAATCTCTGTTTAAATTTTACAGTTAATTAATTAATTAATTAAT 240
DB 242 GATCACTAAGCGGACCTTTCGTCCTCGACGGTGGGCTTTGAGCTCAAGCTCCCTT 301
QY 241 TTATCTTTTATAATAA-----AATAAATACATTTGACACCTCCG 278
DB 302 CTGCTTTTGCACTCGAGGGCAATCTCCGTCGCGCCCGAGGAAACCTTTGACGCGCTCCG 361
QY 279 TTTTATATAGAGGAGACGCGCCAGTCAAACTATCTTATAAATTTGT---TAAAT 335
DB 362 TTACCTTTTGGAGGCGCTACGCGCCCATAGAAACTGTCTACCTGAGACTGTCCCTTGGGCC 421
QY 336 TTGTTTATAAATAATTTATAAGAAATTTATATATATAAATGGTATTTCAATTAACAAT 395
DB 422 GTAAGTCCTGACACAGGTTAGAAATTTCTAGCTCTCCAGAGTGGTATCTCACTGATGGCT 481
QY 396 ACATTATTTCCAAAAAATAATTAATTAATTTCTCCCAATTTATCTGTTATATATATA 455
DB 482 CGGGCCCCCGGAGGAGCGCTTCTTCGCTTCCACCTTAAGCTGGCGAAGAAAGCCCA 541
QY 456 TTTTCAATATCTAATTAATAGTAAGCTTCATAGGCTTCTGCTCTAATATAAAGAAATC 515
DB 542 AAGCCAAATCCAGGAAACAGTGAAGCTTCTAGGCTTCTGCTCAGGTCGAGGTAGTC 601
QY 516 TGCATCTTCACAGATAATTTTATTAAGATTTTATTAAGATTTTATTAAGATTTTATTAAGTCGT 575
DB 602 CGATCTTCACAGACATGTCTAATTTACCGAGCCTCTCTCCGAGACAGATGCGCCAGATCGT 661
QY 576 TACATCTTTTCATGCGAGGTC 594
DB 662 TACGCTTTTCGTCGGGTC 680

RESULT 14
BH537063/c 721 bp DNA linear GSS 14-DEC-2001
LOCUS BH537063
DEFINITION BOGWR91TR BOGW Brassica oleracea genomic clone BOGWR91, DNA
ACCESSION BH537063
VERSION BH537063.1 GI:17772627

KEYWORDS
SOURCE
ORGANISM

GSS.
Brassica oleracea.
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 721)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOGWR91TF
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
Location/Qualifiers

Source
1. .721
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGWR91"
/clone_lib="BOGW"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 163 a 176 c 230 g 152 t
ORIGIN

Query Match 28.5%; Score 169; DB 12; Length 721;
Best Local Similarity 58.6%; Pred. No. 5.9e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACACTACTGCTTACGCTTAGGATGCGA 60
DB 657 GTACCCCTTTAATGGCGAACAGCGCAACCCCTTGGAAACATACACTACAGCCCCAGGTCGCGA 598
QY 61 TAAGCGGACATCGAGTGGCGCAACACCTTTTCGCAATATGAGCTCTCTCGGAAAAGATTAGCC 120
DB 597 AGAGCGGACATCGAGTGGCGCAACACCTTTTCGCAATATGAGCTCTCTCGGAAAAGATTAGCC 538
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGGATAATTTTATTAATTAATTAATTAATTCG 180
DB 537 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGGATAATTTTATTAATTAATTAATTAATTCG 479
QY 181 GATCATTAAGACCGACATTAATCTCTGTTTAAATTTTACAGTTAATTAATTAATTAATTAAT 240
DB 478 GATCACTAAGCGGACCTTTTCGTCCTCGACGGTGGGCTTTGAGTCAAGCTCCCTT 419
QY 241 TTATCTTTTATAATAA-----AATAAATACATTTGACACCTCCG 278
DB 418 CTGCTTTTGCACTCGAGGGCAATCTCCGTCGCGCCCGAGGAAACCTTTTACGCGCTCCG 359
QY 279 TTTTATATAGGAGAGACGCGCCAGTCAAACTATCTCTAATAATATTTGT---TAAAT 335
DB 358 TTACCTTTTGGAGGCGCTACGCGCCCATAGAAACTGTCTACCTGAGACTGTCCCTTGGCCC 299
QY 336 TTTGTTATAAATAATTTTATAAAGAAATTTATATATATAAATGGTATTTCAATTAACAAT 395
DB 298 GTAGGTCCTGACACAAGGTTAGAAATTTCTAGCTCTTCCAGAGTGGTATCTCACTGATGGCT 239
QY 396 ACATTATTTCCAAAAAATAATTAATTAATTTTACTTCCCAATTTATCTGTTATATATATA 455
DB 238 CGGGCCCCCGGAGGAGGCGCTTCTTCGCTTCCACCTTAAGCTGGCGAAGAAAGCCCA 179
QY 456 TTTTCAATATCTAATTAATAGTAAGCTTCATAGGCTTCTTCTCTCTAATATAAAGAAATC 515
DB 178 AAGCCAAATCCAGGAAACAGTGAAGCTTCTATAGGCTTCTTCTGCTGAGGTGCGAGGTAGTC 119
QY 516 TGCATCTTCACAGATAATTTTATTAAGATTTTATTAAGATTTTATTAAGATTTTATTAAGTCGT 575

```

|||||
Db 118 CGCATCTTCACAGACATGCTCTATTTCACCGAGCCTCTCTCCGAGACAGTCCCCAGATCGT 59
|||||
QY 576 TACATCTTTCATGAGGTC 594
|||||
Db 58 TACGCTTTCGTGGGTC 40
|||||

RESULT 15
BH448906/c
LOCUS BH448906 725 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGZI66TR BOG2 Brassica oleracea genomic clone BOGZI66, DNA
sequence.
ACCESSION BH448906
VERSION BH448906.1 GI:17634617
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 725)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOGZI66TF
Contact: Chris Town
TIGR
712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: Sheared ends.
Location/Qualifiers
1..725
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGZI66"
/clone_lib="BOG2"
/notes="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 165 a 178 c 230 g 152 t
ORIGIN

Query Match 28.5%; Score 169; DB 12; Length 725;
Best Local Similarity 58.6%; Pred. No. 5.9e-16;
Matches 363; Conservative 0; Mismatches 230; Indels 26; Gaps 3;

QY 1 GTATCGCTTTAATAGCGAAGACAGACTTACCCCTTAAACATACCTACTGCGCTTAGGATCGGA 60
|||||
Db 636 GTACCGCTTTAATGGCGGACAGCCCAACCCCTTGGAAACATACACAGCCCGCGGTGGCGA 577
|||||
QY 61 TAAGCGACATCGAGGTGCCAAACCTTTTCGTCATATGGACTCTCGGAAAGATTAGCC 120
|||||
Db 576 AGACCGACATCGAGGTGCCAAACCTTCCGTCGATCTGAGCTCTTGGGGAAGATCAGCC 517
|||||
QY 121 TGTATCCCTAGAGTAAGTATTTATCGGTTAAGCGATAATTTTATTAATTAATTAATTCG 180
|||||
Db 516 TGTATCCCTAGAGTAAGTATTTATCCGTTGAGCGAGCGGCCCTTCCACT-CGGCACCGTCG 458
|||||
QY 181 GATCATTAAGACCGACATTAATCTCTGTTTAATTTTAAATTTTACAGTTAATTAATATAT 240
|||||
Db 457 GATCATTAAGCGGACATTTTCGTCGCTCGCGGCTGTCGAGTGGGTCTTGCAGTCAAGCTCCCTT 398
|||||
QY 241 TTATCTTTTATATAA-----AATATAACATTTGTACACCTCG 278
|||||
Db 397 CTGCTTTTGCACTCGAGGGCCAACTCCGTCGCGCCCGAGGAAACCTTTGCAGCGCTCG 338
|||||
QY 279 TTTTATATAGGAGGAGACCGCCCGAGTCAACTATCTTATAAATATTGT---TAAAAAT 335
|||||
```

```

Db 337 TTACCTTTTGGAGGGCTACGCCCCATAGAAACTGTCTAOCCTGAGACTGTCCCTTGGCCC 278
QY 336 TTTGTTTATAAAAAATTTTATAAGAAATTTATATATATATAAATGCTATTTTCAATTAACAATT 395
|||||
Db 277 GTAGGTCCTTGACACACAAGGTTAGAAATTTAGCTCTCTCCAGAGTGGTATATCTCACTATGGCT 218
|||||
QY 396 ACATTATTTCCAAAAAATAATATTACTACTTCCCATTTTATTTCTATGTTATATATATATA 455
|||||
Db 217 CGGGCCCCCGGAGGAGGCGCTTCTTCGCTTCCACCTAAGCTGCGCAGGAAGCCCA 158
QY 456 TTTTCAATATCTATTAATAGATTAAGCTTCATAGGGTCTTTCTGTCCCTAATATAAGAAATC 515
|||||
Db 157 AAGCCAATCCAGGAACAGCTGAAGCTTTCATAGGGTCTTTCTGTCCAGGTGCAGGTAGTC 98
QY 516 TGCATCTTCACAGATAATTTTATTTTCATTAAGATTTTTTTTAAAGACAGCATTTTAAGTCGT 575
Db 97 CGCATCTTCACAGACATGCTCTATTTTACCGAGGCTCTCTCCGAGACAGTGGCCAGATCGT 38
QY 576 TACATCTTTTCATGACAGTC 594
|||||
Db 37 TACGCTTTCGTGGGTC 19
```

Search completed: October 17, 2002, 12:28:14
Job time : 1716 secs

GenCore version 5.1.13
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 10:09:38 ; Search time 48 Seconds
(without alignments)
3039.715 Million cell updates/sec

Title: US-09-369-992c-1_COPY_1147_1740
Perfect score: 594
Sequence: 1 gtagcgtcttaataaggagaa.....ttacatcttttcagcaggtc 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PT05_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	170	28.6	1869	US-08-371-377-21	Sequence 21, Appl
C 2	166	27.9	2904	US-09-463-355-3	Sequence 3, Appl
C 3	154.2	26.0	411529	US-09-103-840A-1	Sequence 1, Appl
C 4	142	23.9	2542	US-09-187-946-3	Sequence 3, Appl
C 5	72.8	12.3	433	US-08-875-972-25	Sequence 25, Appl
C 6	72.8	12.3	16569	US-09-097-889-2	Sequence 2, Appl
C 7	72.8	12.3	16569	US-09-377-856-1	Sequence 1, Appl
C 8	66.4	11.2	639	US-08-998-416-188	Sequence 188, Appl
C 9	65.8	11.1	19124	US-08-487-826B-13	Sequence 13, Appl
C 10	64.6	10.9	629	US-08-454-557C-43	Sequence 43, Appl
C 11	64.6	10.9	629	US-08-340-426D-43	Sequence 43, Appl
C 12	64.6	10.9	629	US-08-450-673C-43	Sequence 43, Appl
C 13	64.6	10.9	629	PCT-US95-17111A-43	Sequence 43, Appl
C 14	63.6	10.7	19124	US-08-487-826B-13	Sequence 13, Appl
C 15	59.2	10.0	665	US-08-883-795A-36	Sequence 36, Appl
C 16	58.4	9.8	3095	5231168-1	Patent No. 5231168
C 17	57.6	9.7	4256	US-08-505-509-31	Sequence 31, Appl
C 18	57.6	9.7	4256	US-08-491-690A-31	Sequence 31, Appl
C 19	57.4	9.7	2061	US-09-187-946-4	Sequence 4, Appl
C 20	57	9.6	6124	US-08-213-419B-3	Sequence 3, Appl
C 21	56.2	9.5	615	US-08-998-416-186	Sequence 186, Appl
C 22	56	9.4	688	US-08-998-416-972	Sequence 972, Appl
C 23	55.2	9.3	8920	US-08-446-855A-1	Sequence 1, Appl
C 24	55.2	9.3	8920	US-09-150-741-1	Sequence 1, Appl
C 25	54.6	9.2	1431	US-09-316-083-2	Sequence 2, Appl
C 26	54.6	9.2	6152	US-08-973-462-1	Sequence 1, Appl
C 27	54.4	9.2	429	US-08-991-789A-215	Sequence 215, Appl

C 28	54.4	9.2	429	4	US-09-062-451-215	Sequence 215, Appl
C 29	54.4	9.2	675	4	US-08-998-416-179	Sequence 179, Appl
C 30	54.4	9.2	676	4	US-08-998-416-280	Sequence 280, Appl
C 31	54.4	9.2	685	4	US-08-998-416-951	Sequence 951, Appl
C 32	54.4	9.2	2058	2	US-08-749-391-1	Sequence 1, Appl
C 33	54.4	9.2	2058	3	US-09-390-200-1	Sequence 1, Appl
C 34	54.2	9.1	665	2	US-08-883-795A-36	Sequence 36, Appl
C 35	54	9.1	837	4	US-08-998-416-288	Sequence 288, Appl
C 36	53.8	9.1	6243	2	US-09-056-075-1	Sequence 1, Appl
C 37	53.6	9.0	615	4	US-08-998-416-186	Sequence 186, Appl
C 38	52.8	8.9	636	4	US-08-998-416-1137	Sequence 1137, Appl
C 39	52.6	8.9	6768	1	US-07-991-755A-1	Sequence 1, Appl
C 40	52.6	8.9	8457	1	US-07-991-867B-1	Sequence 1, Appl
C 41	52.6	8.9	8457	2	US-08-544-332-1	Sequence 1, Appl
C 42	52.4	8.8	2564	1	US-08-224-983-1	Sequence 1, Appl
C 43	52.4	8.8	2564	2	US-08-852-933-1	Sequence 1, Appl
C 44	52.4	8.8	2564	2	US-08-852-945-1	Sequence 1, Appl
C 45	52.4	8.8	2564	2	US-08-853-021-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-371-377-21/c
; Sequence 21, Application US/08371377
; Patent No. 5851764

; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.

; APPLICANT: Shen, Ruqian

; TITLE OF INVENTION: DEVELOPMENT OF DNA PROBES AND

; TITLE OF INVENTION: IMMUNOLOGICAL REAGENTS SPECIFIC FOR CELL SURFACE-EXPRESSED

; TITLE OF INVENTION: MOLECULES AND TRANSFORMATION-ASSOCIATED GENES

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/371,377

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 0575/37590-B

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1869 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: gDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-371-377-21

Query Match

Best Local Similarity 28.6%; Score 170; DB 2; Length 1869;

Matches 375; Conservative 0; Mismatches 215; Indels 26; Gaps 5;

QY 1 GATCGCTTTAATAGCGGAACAGACTTACCCTTAAACATACCTAGCTTAGGATGCCA 60

[illegible]

RESULT 2

```

US-09-465-355/c
Sequence 3, Application US/09465355
Patent No. 6316194
GENERAL INFORMATION:
APPLICANT: Karn, Jonathan
APPLICANT: Knowles, David
APPLICANT: Murchie, Alastair
APPLICANT: Lentzen, Georg
TITLE OF INVENTION: Methods and Kits for Discovery of RNA-Binding Antimicrobials
FILE REFERENCE: 22620/1150 (Formerly 3950/85276)
CURRENT APPLICATION NUMBER: US/09/465.355
CURRENT FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: US 09/325,601
PRIOR FILING DATE: 1999-06-03
PRIOR APPLICATION NUMBER: GB 9812196.5
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: GB 9904790.4
PRIOR FILING DATE: 1999-03-02
PRIOR APPLICATION NUMBER: US 60/122,439
PRIOR FILING DATE: 1999-03-02
PRIOR APPLICATION NUMBER: US 60/088,241
PRIOR FILING DATE: 1998-06-05
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 2904
TYPE: RNA
ORGANISM: Escherichia coli

```

Query Match 27.9%; Score 166; DB 4: Length 2904;
 Best Local Similarity 57.3%; Pred. No. 7e-26;
 Matches 353; Conservative 0; Mismatches 240; Indels 23; Gaps

QY	1	GTATCGCTTTAATAGCGGAACAGACTTACCTTTAAACATACTACTGCCTTAGAGTCGGA	60
Db	2574	GTACCACCTTTAAATGGCAACAGCATACCTTTGGGACCTACTTCAGCCCCAGGATGTGA	2515
QY	61	TAAAGCCGACATCGAGGTGCCAANAACCTTTTCGTCAAATATGGAGCTCTCGGAAAAGATTAGCC	120
Db	2514	TGAGCCGACATCGAGGTGCCAANAACCGCCGTCGATGAATCTTTGGCGGTTATCAGCC	2455
QY	121	TGTTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTAATTAATAATATATCG	180
Db	2454	TGTTATCCCGGAGTAACCTTTTATCCGTTGAGGATGGCCCTTCCATTT-CAGAAACCACCG	2396
QY	181	GATCATTAAAGCCGACATTAAATCTCTGTGTTTAAATTTGTAAATTTTACAGCTTAATATATAT	240
Db	2395	GATCACTATGACCTGCTTTTCGCACCTGCTCGCGCGTCAGCTCGCAGTCAAGCTGGCTT	2336
QY	241	TTATGCTTT-----ATATAATAATATAACATTTGTACACCTCCG	278
Db	2335	ATGCCATTGCACTAACCTCCTGATGTCGACACGAGGATTAGCCAACCTTCGTGCTCCTCCG	2276
QY	279	TTTTTATATAGGAGGAGACCGCCCGAGTCAAACTATCTTATAAATATTGTTAAAAATTTT	338
Db	2275	TTACTCTTTTAGGAGGAGACCGCCCGAGTCAAACTACCCACCGAGACACTGTCCGCAACCCG	2216
QY	339	GTTATAAAAAATTTTAAAGAAATTTATATATATATAATAAATGGTATTTTTCATTAAACAATACA	398
Db	2215	GATTACGGGTCAACGTTAGAACATCAACACATTTAAAGGGTGGTATTTCAAGGTCGGCTCCA	2156
QY	399	TTATTTCCAAAAAATAATATTTACTTTCCCATTTATTTCTATGTTTATATATATATTTT	458
Db	2155	TGCAGACTGGCGTCCACACTTCTAAGCCCTCCCAACCTTATCCTACACATCAAGGCTCAATGT	2096
QY	459	TCAATATCTAATTAAGTAAAGCTTCATAGGCTTTCTGTCCTTATATATAAGAAATCTGC	518
Db	2095	TCAGTGTCAAGCTATAGTAAAGGGTTCACGGGGTCTTTCCGTCTTGCCGGGGTACACTGC	2036
QY	519	ATCTTCAAGATAAATTTTATTCATTAAGATTTTTTTTTTAAGCAGCATTTTAAGTCGTTAC	578
Db	2035	ATCTTCAAGCGAGTTCAATTTCACTGAGTCTCGGGTGGAGACAGCGCTGCCCATCATTAC	1976
QY	579	ATCTTTTCATCGAGTTC	594
Db	1975	GCATTCGTGCGAGTTC	1960

RESULT 3

```

US-09-103-840A-1/c
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103.840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 441529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	26.0%	Score 154.2;	DB 4;	Length 4411529;
Best Local Similarity	58.2%	Pred. No. 3.7e-23;		
Matches	359;	Conservative 0;	Mismatches 233;	Indels 25; Gaps 4;
QY 1	GTATCGCTTTAATAGGCGAACAGACTTACCTTAAACATACACTACTGCCCTAGGATCGGA 60			
Db 1476466	GTGCCCGCTTTATGGCGGAACACGCCAACCCCTTGGGAGCTGCTCCAGCCCGGAGATGCGGA 1476407			
QY 61	TAAGCGGACATCGAGGTGCCAAACGCTTTCTCGAATATGGCACTTCGGAAGAAGATTAGCC 120			
Db 1476406	CGAGCGCGACATCGAGGTGCCAAACCACTCCGTCGATATGGACTCTTGGGGAAGATCAGCC 1476347			
QY 121	TGTTATCCCTAGAGTAAGCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180			
Db 1476346	TGTTATCCCGGGTACCTTTTATCCGTTGAGGAGCACCCCTTCCACT -CGGGGGTGCCG 1476288			
QY 181	GATCATTTAAGACCGACATTAATCTCTGTTAAATTTGTAATTTTACAGTTAAATTTATATAT 240			
Db 1476287	GATCACTAATCCGACTTTTCGTCCCTGCTTGACTTGTAAAGTCTCGCAGTCAAGCTCCCTT 1476228			
QY 241	TTATCTTTA -----TATAATAATATAACATTTGTACACCTCCG 278			
Db 1476227	GTGCATTTTACATCGCCACCTGATTCGCGTTCAGGTTGAGGGAACCTTTTGGCGCCTCCG 1476168			
QY 279	TTTTTATATAGGAGGACCGCCCGAGTCAACATCTATTATAAATATGTTTAAAAATTTT 338			
Db 1476167	TTACATTTTTAGGAGCAACGCCCCAGTTAACTACCCCGCAGGCACTGTCCCTAAACCC 1476108			
QY 339	GTTATAAAAT -TTTATAAGAATTTATATATATATAATAAATGGTATTTCAATAACAATTAC 397			
Db 1476107	GATTCAGGGTTTCAGGTTTAGATGCCCAATACGATCAGATGGTATTTTCAACACGACTCC 1476048			
QY 398	ATTATTTCCAAAAAATAATTACTACTTCCCATTTATTTCTATGTTATATATATATATT 457			
Db 1476047	GCCCAACTGGCGTCGAGGTTTCACAGTCTCCACCTATCTTACACAACCGTACCGAAC 1475988			
QY 458	TTCAATATCTATTAAATAGTAAAGCTTCATAGGTCCTTCTGTCCTCTAAATATAAGAAATCTG 517			
Db 1475987	ATCAATATACCAAGTTGTAGTGAAG -GTCCCGGGGCTCTTTTCGTCCTGCCCGGGTAACGAG 1475929			
QY 518	CATCTTCCACAGATAAATTTTATTCATTAAGATTTTTTTTAAGCAGCATTTTAAGTCGTTA 577			
Db 1475928	CATCTTTACTCGTAGTGCATTTTCGCCGAGTCTATCGTTGAGACAGTTGAGAAGTCGTTA 1475869			
QY 578	CATCTTTTCATGCAGGTC 594			
Db 1475868	CGCATTTCTGTCGAGGTC 1475852			

```

RESULT 4
US-09-187-946-3/c
Sequence 3, Application US/09187946
Patent No. 6255467
GENERAL INFORMATION:
APPLICANT: Lindner, Luther E.
APPLICANT: MacPhee, Kathleen
TITLE OF INVENTION: Human Blood Bacteria
FILE REFERENCE: D6026
CURRENT APPLICATION NUMBER: US/09/187-946-3
CURRENT FILING DATE: 1998-11-02
EARLIER APPLICATION NUMBER: US 60/091879
EARLIER FILING DATE: 1997-11-06
NUMBER OF SEQ ID NOS: 20
SEQ ID NO 3
LENGTH: 2542
TYPE: DNA
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: Rb 23S rRNA sequence
US-09-187-946-3

```

Query Match

23.98; Score 142; DB 4; Length 2542;

	Best Local Similarity	54.9%;	Pred. No.	5.2e-21;		Gaps	2;	
	Matches	338;	Conservative	0;	Mismatches	255;	Indels	23;
QY	1	GTATCGCTTTTAANTAGGCGAACACAGACTTACCCTTAAAAACATACTACTGCCCTTAGGATCGGA	60					
DB	2373	GTACCACTTTAATCGGGCAGACGCAGCAACCTTGCGACCTTCGACCCCGAGGATGTGA	2314					
QY	61	TAAGCCGACATCGAGGTGCGAAACGCTTTTCGTCAANATGGAGCTCTCGGAAGAAGATTWACC	120					
DB	2313	TGAGCCGACATCGAGGTGCCAAACGACCGCGCATATGGAGCTTTGGGGGTCTATCACGCC	2254					
QY	121	TGTTATCCCCTAGAGTAAGTCTTTTATCCGTAAAGCGATAATTTTATATTAATTAATTATCG	180					
DB	2253	TGTTATCCCCGGGTACTTTTTATCCGTGTAGCGAT--GGCCCCACCGCGGGGACCACCG	2195					
QY	181	GATCATTTAAGACCCGACAATTAATCTCTCTTTAAATTTTAAAATTTTACAGTTAAATATATAT	240					
DB	2194	GATCACTATGACGACTTCGTCTCTCTCGACATGTCGCTCGCAGTCAAGCGGCGCTT	2135					
QY	241	TTATCTTTTATA-----TAAATAATATAACATTTGTACACCTCCG	278					
DB	2134	ATGCGATTGGACGGGACGAGCGATTTC CGACCGCTCTGAGCCCCACCCTTCGTAGCGCTCCG	2075					
QY	279	TTTTTATATAGGAGGACCGCCCCAGCTCAAACATCTTATAAATATTTGTAAATAATTTT	338					
DB	2074	TTAGCGTTTGGAGGCGACCGCCCCAGCTCAAACATGCGCTGCGCGGTCCCGCGCGGCC	2015					
QY	339	GTTATAAAAAATTTTATAAGAAATTTATATATATAFAAAATGGTATTTCATTAAACAATTACA	398					
DB	2014	GATCAGGCGCGCGGTTAGACCACCATATCGTCAAGSGTGGTATTTC AAGGTGGCTCCA	1955					
QY	399	TTATTTCCAAAAANAATATTACTCTCCCATTTATCTATGTATTATATATATATATTT	458					
DB	1954	TCCAGGCTGGCGGCCGGAACITCAAAGGCTACCACCTATCCTACACATGCGGACACGGAAG	1895					
QY	459	TCAATATCTATTATAGTAAGCTTTCATAGGCTTTCTCTCTCTTAATATAAGAAATCTGC	518					
DB	1894	CCAGCGCAAAGCTACAGTAAGGTGACGGGGTCTTTCCCTGTGACCGAGGAACCCCGC	1835					
QY	519	ATCTTCACAGATAATTTTATTTTCATTTAAGATTTTTTTTAAGACAGCATTTTAAGCTGTAC	578					
DB	1834	ATCTTCAGGGGAATCAATTTTCAC TGAGCGGATGTGGAGACAGCGGGGAGATCGTTAC	1775					
QY	579	ATCTTTTCATCGAGTC	594					
DB	1774	GCCATTCGTGAGGTC	1759					

RESULT 5
US-08-875-972-25/c
; Sequence 25, Application US/08875972
; Patent No. 5985564
; GENERAL INFORMATION:
; APPLICANT: Huntington Potter and Jinhue Li
; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
; TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,972
; FILING DATE: 08-AUG-97
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,448
; FILING DATE: 16-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan Esq., Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-03PA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-875-972-25

Query Match 12.3%; Score 72.8; DB 2; Length 433;
Best Local Similarity 62.8%; Pred. No. 5.1e-07;
Matches 113; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 1 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATACCTGCTGCGAAAGATTAGCC 60
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 GTAGGACTTTAATCGTTGAACAAACGAAACCTTTAATAGCGGCTGCACCATCGGGATGCC 199

Qy 61 TAAGCCACATCGAGGTCGCAACACCTTTTCGTCATATGACTCTCGGAAAGATTAGCC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TGATCCAACTCGAGGTCGTAACCCCTATTGTTGATAGACTCTAGATAGGATTGCCG 139

Qy 121 TGTATCCCTAGACTACTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 TGTATCCCTAGGTAACCTTTGTCGTTGTCGTAAGTTATGGATCAATTCAGTATAGTAG 79

RESULT 6
US-09-897-889-2/c
; Sequence 2, Application US/09097889
; Patent No. 621817
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 16569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-097-889-2

Query Match 12.3%; Score 72.8; DB 4; Length 16569;
Best Local Similarity 62.8%; Pred. No. 7.3e-07;
Matches 113; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 1 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATACCTGCTGCGAAAGATTAGCC 60
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3060 GTAGGACTTTAATCGTTGAACAAACGAAACCTTTAATAGCGGCTGCACCATCGGGATGCC 3001

Qy 61 TAAGCGGACATCGAGGTCGCAACACCTTTTCGTCATATGACTCTCGGAAAGATTAGCC 120
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3000 TGATCCAACTCGAGGTCGTAACCCCTATTGTTGATAGACTCTAGATAGGATTGCCG 2941

Qy 121 TGTATCCCTAGACTACTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2940 TGTATCCCTAGGTAACCTTTGTCGTTGTCGTAAGTTATGGATCAATTCAGTATAGTAG 2881

RESULT 7
US-09-377-856-1/c
; Sequence 1, Application US/09377856
; Patent No. 6344322
; GENERAL INFORMATION:
; APPLICANT: Polyak, Kornelia
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Subtle Mitochondrial Mutations as Tumor
; TITLE OF INVENTION: Markers
; FILE REFERENCE: 1107-82346
; CURRENT APPLICATION NUMBER: US/09/377,856
; CURRENT FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/097,307
; PRIOR FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 16569
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-377-856-1

Query Match 12.3%; Score 72.8; DB 4; Length 16569;
Best Local Similarity 62.8%; Pred. No. 7.3e-07;
Matches 113; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 1 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATACCTGCTGCGAAAGATTAGCC 60
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3060 GTAGGACTTTAATCGTTGAACAAACGAAACCTTTAATAGCGGCTGCACCATCGGGATGCC 3001

Qy 61 TAAGCGGACATCGAGGTCGCAACACCTTTTCGTCATATGACTCTCGGAAAGATTAGCC 120
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3000 TGATCCAACTCGAGGTCGTAACCCCTATTGTTGATAGACTCTAGATAGGATTGCCG 2941

Qy 121 TGTATCCCTAGACTACTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2940 TGTATCCCTAGGTAACCTTTGTCGTTGTCGTAAGTTATGGATCAATTCAGTATAGTAG 2881

RESULT 8
US-08-998-416-188
; Sequence 188, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippssen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jorgen
```

APPLICANT: Knechtle, Philipp
 APPLICANT: Reibischung, Corinne
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 1152
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 6239264artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: No. 6239264th Carolina
 COUNTRY: USA
 ZIP: 27709

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998,416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689

SEQUENCE CHARACTERISTICS:
 INFORMATION FOR SEQ ID NO: 188:
 LENGTH: 639 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: PAG1075RP

US-08-998-416-188

Query Match 11.2%; Score 66.4; DB 4; Length 639;
 Best Local Similarity 68.9%; Pred. No. 1.1e-05;
 Matches 91; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 463 TATCTATTATAGTAAGCTTCATAGGGCTTTCTGCTCCTAATATAGAAATCTGCATCT 522

Db 5 TATCTAATTACAGTAAGCTGCAAGGGCTTTTCTGCTCTTCTACAAATACCTAGCATCT 64

QY 523 TCACAGATAATTTTATTTCATTAAGATTTTTTTTAAAGACAGCATTTTAAGTCGTTACATCT 582

Db 65 TCACTAAGATTTCAATTCACCTAGATTAAAGGAGAGACAGTTGTTGTATCATCTACGTC 124

QY 583 TTCATGCAGGTC 594

Db 125 TTCATGCAGGAC 136

RESULT 9
 US-08-487-826B-13
 Sequence 13, Application US/08487826B
 Patent No. 5993827
 GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhaun
 APPLICANT: Wellem, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,826B
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Israelsen, Ned
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NIH121.001CPI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176

SEQUENCE CHARACTERISTICS:
 INFORMATION FOR SEQ ID NO: 13:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

Query Match 11.1%; Score 65.8; DB 2; Length 19124;
 Best Local Similarity 51.6%; Pred. No. 2e-05;
 Matches 181; Conservative 0; Mismatches 162; Indels 8; Gaps 1;

QY 137 ACTTTTATCGTTAAGCGATAATTTTATTATAAATAATTCGATCATTAAGACCGAC 196

Db 6773 ACTTTTATATGTTGTAATTTTCGTAATTTTTCATTTTAAATTTTACTTAATA 6832

QY 197 ATTAATCTCTGTTTAAATTTTACAGTAAATATATATATATATATATATATATAT 256

Db 6833 AATAAACAATAAAAAAATAATATATATATATATATATATATATATATATATAT 6892

QY 257 AATAAACAATGTACACCTCCGTTTATATAGGAGAGACCCGCCAGTCACACTATCT 316

Db 6893 AATAAATAATTTCTGATATATATATATATATATATATATATATATATATATAT 6952

QY 317 TATAAATAATTTGTTAAAAATTTTGTATAAAAAATTTTATAAGAAAT-----TTATATAT 368

Db 6953 ATTAAT 7012

QY 369 ATATAAAATGGTATTTTCAATTAACAATTTACATATTTTCCAAAAAATAATTTACTACTTC 428

Db 7013 ATATTAATAAT 7072

QY 429 CCATTTTCTATGTTATATATATATATATATTTTCAATATCTATTAATAGTAAA 479

Db 7073 TAAT 7123

RESULT 10
 US-08-454-557C-43
 Sequence 43, Application US/08454557C
 Patent No. 5830670
 GENERAL INFORMATION:
 APPLICANT: de la Monte, Suzanne
 APPLICANT: Wands, Jack R.
 TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
 TITLE OF INVENTION: of Alzheimer's Disease
 NUMBER OF SEQUENCES: 121

Query Match 10.9%; Score 64.6; DB 2; Length 629;
Best Local Similarity 64.2%; Pred. NO. 2.5e-05;
Matches 97; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

18	QY	GACAGACCTTACCCCTTAACACATCACTACTGCTTACGATCGGATAGCCGACATCGAGGT	77
206	Db	GACAAACGACACCTTTAATAGCGGCTGCACCATCGGATGCTCTGATCAACATCGAGGT	265
78	QY	GCCAAACCTTTTCGTCAAATAGACGCTCTCGGAAAGATTAGCCTGTTATCCCTAGAGTAA	137
266	Db	CGTAACCCCTATTGTTGATAGACGCTCAAGATAGGATTGCGCTGTTATCCCTAGGTTAA	325
138	QY	CTTTTATTCGCTTAAGCGATAATTTTATTATT	168
326	Db	CTTGTTCCCGTTGGTCAAGTTATTGGATCAAT	356

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 629 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-43

Query Match	10.9%	Score 64.6;	DB 5;	Length 629;
Best Local Similarity	64.2%;	Pred. No. 2.5e-05;		
Matches 97;	Conservative	0;	Mismatches 54;	Indels 0;
				Gaps

Qy	18	GAACACATTACCTTTAAACATCTACTTCCTTAGGATCGCATAGCCGACATCGAGT	77
Db	206	GAACAAAGCACTTTAATAGCGCTGCACCATCGGAGTCTCTGATCCAACATCGAGT	265
Qy	78	GCCAAACCTTTTCGTCAATATGAGCTCTCGAAAAGATTAGCTGTGTATCCCTCAGAGTAA	137
Db	266	CGTAAACCCCTATTGTGATATGAGCTCTAGAATAGGATTGCGCTGTTATCCCTCAGGTTAA	325

QY 138 CTTTATCCGTTAAGCGATAATTATTAAT 168
||| | - - - ||| - - -
Db 326 CTTGTCCGTTGGTCAAGTTATTCGATCAAT 356

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match	10.7%	Score 63.6;	DB 2;	Length 19124;
Best Local Similarity	52.5%;	Pred. No. 5.5e-05;		
Matches 166;	Conservative	0;	Mismatches 144;	Indels 6;
				Gaps 1;

DQ

F06879

Cy 158 ATTTTATTATAAATTAATTATCGGATCATTAAGACGGACATTAACTCTGTGTTTTAATTCCT 217
||||| |||| | |||| |||| | |||| |||| | |||| | |||| |
Bb 15773 ATTTTTTTAAAATTTTTTTTTTATTTATC GATATATATTTTATTTTAAATATATTTTTTTCCCT 15714

Qy 218 AAATTTACAGTAAATATATATATTATCTTTATATAAATAACATGTACACCCTC 277
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15713 TTTTTTTGTTTTATCATATATATTTTTTTTTTTTAATGTTTTTTTTTTCTCTCTTTT 15654

Qy 278 GTTTTATATAGGAGGACGCCCCCGTCAAACTATCTATAAATATGTAAAAATT 337
||||||| | | | ||||| | | |
Dd 15653 GTTTTATATTTTTTAATC-----ATTTTTTTTTTATATAAAAATTTTTTTTAAATT 15600

Qy 338 TGGTATAAAAAATTTTATAAGAAATTTATATATATATAAATGGTATTTTCATTACAATTAC 397
| | | | | | | | | | | | | | | | | |
Db 15599 TTTTTGATAATCTTTTTTCATTTTATTTCTCATCAAAATTTATATTTTATTATAAATTTT 15540

Qy	398	ATTATTTTCCAAAAAATAATATTACTACTTCCCAATTTATCTCATGTATATATATATAT	457
Db	15539	ATTATTTTAAAAAATTTTCTCCTTTTTTTTTTTTTTTTATTTTAAATAAATTT	15480
Qy	458	TTCAATATCTATTAAT	473
Db	15479	TTTTTATATTTCAAT	15464

RESULT 15

```

US-08-883-795A-36/C
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Michelle
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: Rh 32
US-08-883-795A-36

```

	Query Match	10.0%;	Score 59.2;	DB 2;	Length 665;
	Best Local Similarity	48.3%;	Pred. No. 0.00031;		
	Matches 195;	Conservative 0;	Mismatches 208;	Indels 1;	Gaps 1;
QY	155	ATAAATTTTATTTAAATAAATTCGGATCATTAAGACCGACACATTAATCTCTGTTTAATT	214		
Db	485	AAATATGTAATTTATAAATACTTTATAAAATATGTAATTTATAAAATATGTAATTTATAAACA	426		
QY	215	TGTAATTTTACAGTTAATATATATTTATCTTTTATATAATAATATAACATGTGACACC	274		
Db	425	TTTTAATATAAATATGTAATTTATAAACAATTTTAATTTATAAAATATGTAATTTATAAACA	366		
QY	275	TCGGTTTTTATATAGGAGGACCGCCCGAGTCAAATCTCTTATATAATTTGTTAAAAA	334		
Db	365	TTTTAATATAAATATGTAATTTATAAACAATTTTAATTTATAAAATATGTAATTTATAAACA	306		
QY	335	TTTTGTGTTATAAAATTTTATAAGAAATTTATATATATATAAAATGGTATTTCAATTAACAT	394		

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 10:04:08 ; Search time 217 Seconds
(without alignments)
4699.757 Million cell updates/sec

Title: US-09-369-992C-l_COPY_1147_1740
Perfect score: 594
Sequence: 1 gtatcggttaatagcgaa.....ttacatatttcagcaggtc 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	594	100.0	5849	19 AAV331135	Plasmodium berghei
2	204	34.3	3398	20 AAX20282	Borrelia burgdorferi
3	204	34.3	116277	20 AAX20249	Borrelia burgdorferi
4	190	32.0	1385	22 AAH54997	S. epidermidis gen
5	190	32.0	2725	22 AAH54148	S. epidermidis gen
6	190	32.0	2869	22 AAH55038	S. epidermidis gen
7	190	32.0	3012	22 AAH54319	S. epidermidis gen
8	190	32.0	3012	22 AAH54343	S. epidermidis gen
9	190	32.0	3030	22 AAH54552	S. epidermidis gen

10	190	32.0	3086	22 AAH54819	S. epidermidis gen
11	190	32.0	3099	22 AAH54513	S. epidermidis gen
12	190	32.0	3229	22 AAH54504	S. epidermidis gen
13	190	32.0	3232	22 AAH54853	S. epidermidis gen
14	190	32.0	3405	22 AAH54539	S. epidermidis gen
15	190	32.0	3500	22 AAH54518	S. epidermidis gen
16	190	32.0	3513	22 AAH53987	S. epidermidis gen
17	190	32.0	3608	22 AAH54395	S. epidermidis gen
18	190	32.0	3625	22 AAH54294	S. epidermidis gen
19	190	32.0	3656	22 AAH54375	S. epidermidis gen
20	190	32.0	3716	22 AAH54078	S. epidermidis gen
21	190	32.0	3845	22 AAH54156	S. epidermidis gen
22	190	32.0	3929	22 AAH54246	S. epidermidis gen
23	190	32.0	3937	22 AAH54408	S. epidermidis gen
24	190	32.0	4106	22 AAH54320	S. epidermidis gen
25	188.4	31.7	892	23 AAS51075	Staphylococcus aur
26	188.4	31.7	2923	23 AAC89402	Staphylococcus aur
27	188.4	31.7	5048	21 AAA65735	Streptococcus pneu
28	188.4	31.7	5273	20 AAX24982	Haemophilus influe
29	188.4	31.7	5519	20 AAX24981	Haemophilus influe
30	188.4	31.7	6591	18 AAV77425	Staphylococcus aur
31	188.4	31.7	1830121	17 AAT42063	Haemophilus influe
32	185.2	31.2	269223	22 AAF28554	Genomic fragment #
33	178	30.0	2405	20 AAX13746	Enterococcus faeca
34	175.8	29.6	2922	20 AAX03532	Chlamydia trachoma
35	174	29.3	2904	21 AAA66047	E. coli proliferat
36	174	29.3	2904	21 AAA66052	E. coli proliferat
37	174	29.3	2904	22 AAF23016	E. coli 23S rRNA s
38	174	29.3	2904	22 AAC89403	Sequences from 23S
39	174	29.3	2907	19 AAV38096	Enterohaemorrhagic
40	174	29.3	2907	19 AAV38107	Enterohaemorrhagic
41	174	29.3	3118	22 AAH49806	Escherichia coli t
42	174	29.3	5097	20 AAX24983	E. coli MG1655 rrr
43	174	29.3	5105	20 AAX24989	E. coli MG1655 rrr
44	172.4	29.0	5090	20 AAX24988	E. coli MG1655 rrr
45	172.4	29.0	30246	18 AAV74367	Staphylococcus aur

ALIGNMENTS

RESULT 1
AAV33135
ID AAV33135 standard; DNA; 5849 BP.
XX
AC AAV33135;
XX
DT 07-DEC-1998 (first entry)
XX
DE Plasmodium berghei plastid PSI-PL470 gene.
XX
KW Malaria; infection; therapy; diagnosis; vaccine; plastid;
KW PSI-PL470 gene; ds.
XX
OS Plasmodium berghei ANKA strain.
XX
PN WO9835057-Al.
XX
PD 13-AUG-1998.
XX
PF 05-FEB-1998; 98WO-IB00212.
XX
PR 26-SEP-1997; 97AU-0009481.
PR 06-FEB-1997; 97AU-0004953.
PR 21-APR-1997; 97AU-0006329.
XX
PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Kara AKU, Nelson JS, Tan TMC, Tham JM, Ting RCY;
XX WPI; 1998-447251/38.

PT Detecting Plasmodium infection from hybridisation with
PT extrachromosomal element - providing genus or species specific
PT diagnosis with few false negatives, in humans or animals

XX Claim 15; Page 54-59; 120pp; English.

XX This is the nucleotide sequence of one strand of the PSI-PL470
CC gene of the 30.7 kb extrachromosomal plasmid of Plasmodium berghel.
CC This plasmid encodes organelle-like rRNAs, tRNAs, ribosomal
CC proteins and RNA polymerase subunits, amongst others. Plasmodium is
CC detected in a human or animal sample by treating it, or derived
CC nucleic acid, with a Plasmodium extrachromosomal genetic element or
CC derived nucleic acid (A) and detecting any hybridisation. (A) can
CC include the PSI-PL470, PLH-PPH, PRB or PWQ gene, the mitochondrial
CC COXI gene, and nucleic acids derived from them. Also new are
CC (A)-specific probes and primers (see AAV33139-56). The method is
CC used to diagnose Plasmodium infection. Also (not claimed) the
CC polypeptides encoded by (A) are useful as targets for drug
CC development and for development of anti-malaria vaccines. The high
CC degree of similarity between (A) from different species allows
CC development of genus- or species-specific assays that result in
CC fewer false negatives than known methods (typically 1% against 3%).

XX Sequence 5849 BP; 2296 A; 673 C; 557 G; 2323 T; 0 other;

Query Match 100.0%; Score 594; DB 19; Length 5849;
Best Local Similarity 100.0%; Pred. No. 3.5e-79;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTATCGCTTTAATAGGCGAACAGACTTACCGTTAAACATACACTACTCGCTTAGATCGGA 60

Db 1147 GTATCGCTTTAATAGGCGAACAGACTTACCGTTAAACATACACTACTCGCTTAGATCGGA 1206

QY 61 TAAGCGGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAAAGATTAGCC 120

Db 1207 TAAGCGGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAAAGATTAGCC 1266

QY 121 TGTATCCCTAGAGTAAGCTTTTATCCGTTAAGCGATAATTTTATTTAATAATTAATATCG 180

Db 1267 TGTATCCCTAGAGTAAGCTTTTATCCGTTAAGCGATAATTTTATTTAATAATTAATATCG 1326

QY 181 GATCATTAAGACCGACATTAATCTCTGTTAAATTTTACAGTTAATATATAT 240

Db 1327 GATCATTAAGACCGACATTAATCTCTGTTAAATTTTACAGTTAATATATAT 1386

QY 241 TTATCTTTATATAATAATATAACATTTGACACCTCGCTTTTATATAGGAGGAGCCGC 300

Db 1387 TTATCTTTATATAATAATATAACATTTGACACCTCGCTTTTATATAGGAGGAGCCGC 1446

QY 301 CCCAGTCAAACTATCTTATAAATATTTGTTAAAAATTTTGTATAAAAAATTTTATAAGAA 360

Db 1447 CCCAGTCAAACTATCTTATAAATATTTGTTAAAAATTTTGTATAAAAAATTTTATAAGAA 1506

QY 361 TTATATATATAATAAATGGATTTCATTAACAATACATATTTCCTCAAAAAATATATT 420

Db 1507 TTATATATATAATAAATGGATTTCATTAACAATACATATTTCCTCAAAAAATATATT 1566

QY 421 ACTACTTCCCATTTATCTATGTTATATATATATATTTTCAATATCTATTAAATAGTAAAG 480

Db 1567 ACTACTTCCCATTTATCTATGTTATATATATATATTTTCAATATCTATTAAATAGTAAAG 1626

QY 481 CTTCATAGGGTCTTCTGCTTAATAAAGAAATCTGCATCTTTCACAGATAATTTTATTT 540

Db 1627 CTTCATAGGGTCTTCTGCTTAATAAAGAAATCTGCATCTTTCACAGATAATTTTATTT 1686

QY 541 CATTAAGATTTTTTTTAAGACAGCATTTAAGTCGTTACATCTTTTCATCGAGGTC 594

Db 1687 CATTAAGATTTTTTTTAAGACAGCATTTAAGTCGTTACATCTTTTCATCGAGGTC 1740

RESULT 2

AXX20282/c

ID AAX20282 standard; DNA; 3398 BP.

XX

AXX20282;

AC

DT 04-MAY-1999 (first entry)

XX

Borrelia burgdorferi polynucleotide sequence #35.

XX

Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
infection; diagnosis; characterisation; detection; ds.

OS

Borrelia burgdorferi.

XX

WO9858943-A1.

XX

30-DEC-1998.

XX

18-JUN-1998; 98WO-US12764.

XX

03-SEP-1997; 97US-0057483.

PR

20-JUN-1997; 97US-0050359.

PR

22-JUL-1997; 97US-0053344.

PR

22-JUL-1997; 97US-0053377.

XX

(HUMA-) HUMAN GENOME SCI INC.

PA

(MEDI-) MEDIMMUNE INC.

XX

Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;

PI

White OR;

XX

WPI; 1999-081217/07.

XX

New isolated Borrelia burgdorferi nucleic acids - used to develop

PT

products for the detection, diagnosis, characterisation, prevention
and therapy of infections, particularly Lyme disease

PT

XX

Claim 1; Page 998-1000; 1128pp; English.

PS

Sequence 3398 BP; 1096 A; 535 C; 869 G; 896 T; 2 other;

XX

Query Match 34.3%; Score 204; DB 20; Length 3398;

Best Local Similarity 62.5%; Pred. No. 4.7e-22;

Matches 385; Conservative 0; Mismatches 200; Indels 31; Gaps 3;

QY 1 GTATCGCTTTAATAGGCGAACAGACTTACCGTTAAACATACACTACTGCTTAGATCGGA 60

Db 2995 GTACCGCTTTAAATGGCGAACAGCCATACCGTTAGGACCTGCTCCAGCCCTAGGATGCGA 2936

QY 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAAAGATTAGCC 120

Db 2935 TGAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAAAGATTAGCC 2876

QY 121 TGTATCCCTAGAGTAAGCTTTTATCCGTTAAGCGATAATTTTATTTAATAATTAATATCG 180

Db 2875 TGTATCCCTAGAGTAAGCTTTTATCCGTTAAGCGATAATTTTATTTAATAATTAATATCG 2817

QY 181 GATCATTAAGACCGACATTAATCTCTGTTAAATTTTACAGTTAATATATATAT 240

Db 2816 GATCATTAAGACCGACATTAATCTCTGTTAAATTTTACAGTTAATATATATAT 2757

QY 241 TTATCTTTA-----TATAATAATAATAACATTTTACACCTCG 278

Db 2756 ATGCGCTTTACACTTACAGATGATTTCCCAACCACTTAAGGTAACCTTTGGCACCTCCG 2697

Db 1421 ATCTTCACAGGTACTATGATTTCACCGAGTCTCTCGTTGAGACAGTCCCAAAATCGTTAC 1480
QY 579 ATCTTTTCATGCGAGGTC 594
Db 1481 GCCTTTTCGCGGGTC 1496
RESULT 7
AAH54319/C
ID AAH54319 standard; DNA; 3012 BP.
AC AAH54319;
DT 03-SEP-2001 (first entry)
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3683.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
XX WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
XX WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1293-1294; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 3012 BP; 842 A; 612 C; 849 G; 709 T; 0 other;
Query Match 32.0%; Score 190; DB 22; Length 3012;
Best Local Similarity 59.7%; Pred. No. 5.3e-20;
Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;
QY 1 GTATCGCTTTAATAGCGGACACAGCTTACCCCTTAAACATCACTACTGCGCTTAGGATCGCA 60
Db 2261 GTACCGCTTTAATAGCGGACACAGCCCAACCCCTTGGGACCCACTACAGCCCAAGATCGCA 2202

QY 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAAAGATAGCC 120
Db 2201 TGAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGGACTCTCGGAGATAGCC 2142
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAAATTTATATTAATAATATCG 180
Db 2141 TGTATCCCGGGGTAGCTTTTATCCGTTGAGCGATGGCCCT-TCCATGCGGAACACCG 2083
QY 181 GATCATTAAAGACCGACATTAATCTCTCTTTAAATTTTAAATTTTACAGTTATATATAT 240
Db 2082 GATCATTAAAGACCGACATTAATCTCTCTTTAAATTTTAAATTTTACAGTTATATATAT 2023
QY 241 TTATCTTTATATAAATAAT- - - - - AACATTTGACACCTCCG 278
Db 2022 ATGCTTTTACACTCTATGAATGATTTCCAAACCTTCTGAGGGAACCTTTGAGCGCTCCG 1963
QY 279 TTTTATATAGGAGGACCGCCCGCCAGTCAAACTATCTTATAAATATTTTAAAAATTTT 338
Db 1962 TTACCTTTTAGAGGCGGACCGCCCGCCAGTCAAACTGCGCGCTGACACTGTCTCCCAAC 1903
QY 339 GTTATAAAATTTTATAAGAAATTTATATATATAATAAATGATTTTCAATTAACAATATACA 398
Db 1902 GATAAGTGGTGGGTTAGAAAGCCACACAGCTAGGGTAGTATCCCAACACGCTCCA 1843
QY 399 TTATTTCCAAAAAATAATATTAATCTACTTCCCAATTTATCTATGTTATATATATATATTT 458
Db 1842 CGTAAGCTAGCGCTCAGCTTTGAAAGGCTCCTACCTATCCTGTACAAGCTGTGCGGAAT 1783
QY 459 TCAATATCTAATAAGCTTCAAGGCTTCAAGGCTCTTTCTGTCCTTAATAAAGAATCTGC 518
Db 1782 TCAATATCAGGCTACAGTAAAGCTCCACGGGCTCTTCCGCTCGCGGGTAACCTGC 1723
QY 519 ATCTTCACAGATAATTTATTTTCAATTAAGATTTTTTTTAAAGACAGCATTTAAGTCGTTAC 578
Db 1722 ATCTTCACAGTACTATGATTTTCAACCGAGTCTCTCGTTGAGACAGTCCCAAAATCGTTAC 1663
QY 579 ATCTTTTCATGCGGTC 594
Db 1662 GCCTTTTCGTCGCGGTC 1647
RESULT 8
AAH54343
ID AAH54343 standard; DNA; 3012 BP.
AC AAH54343;
XX
XX 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3707.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX

Claim 8; Page 1323-1324; 2188pp; English.

Claim 8; Page 1323-1324; 2180pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3012 BP: 692 A; 887 C; 653 G; 780 T; 0 other:

Query Match 32.0%; Score 190; DB 22; Length 3012;
Best Local Similarity 59.7%; Pred. No. 5.3e-20;
Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;

Qy	1	GTATCGCTTTAATAGCGGAACAGACTTACCCCTTAAACATACTACTGCTTTAGGATCGGA	60
Db	942	GTACCGCTTTAATGGCGAACAGCCCAACCTTGGGACCGGACTACAGCCCCAGAGATCGGA	1001
Qy	61	TAAGCGGACATCGAGGTGCGAAACCTTTTCGTCAATATGGACTCTCGGAAAAGATTAGCC	120
Db	1002	TGAGCGGACATCGAGGTGCGAAACCTTCCCGTCGATGTGAACCTCTTGGGGAGATAGACC	1061
Qy	121	TGTTATCCCTAGTAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAAATATTATCG	180
Db	1062	TGTTATCCCGGGGTAGCTTTTATCCGTTGAGCGATGGCCCT-TCCATGCGGAACCAACCG	1120
Qy	181	GATCATTAAAGACCGGACATTAATCTCTGTTTAAATTGTTAAATTTTACAGTTAAATTATAT	240
Db	1121	GATCATAAGTCGCTCTTTTCGACCCCTGCTCGACTTGTAGGTCTCGCAATCAAGCTCCCTT	1180
Qy	241	TTATCTTTTATATAAATAATAT-----AACATTGTACACCTCCG	278
Db	1181	ATGCCCTTACACTCATGAATGATTCCAACCATCTTCAGGGAACCTTGGCGCCTCCG	1240
Qy	279	TTTTTATATAGGAGGACGCCGCCAGTCAACTATCTTATAAATATGTTTAAAAATTTT	338
Db	1241	TTACCTTTTAGGAGGCGACGCCGCCAGTCAAACTGCCCGCTGACACTGTCTCCACACAC	1300
Qy	339	GTATATAAAATTTTATAAGAAATTTATATATATATAAAATGCTATTCATTAAACAATTACA	398
Db	1301	GATAGTGGTGGGGTTAGAAAGCCACACAGCTAGGGTAGTAGTCCACCACCAACGCCCTCCA	1360
Qy	399	TTATTTCAAAAAATAATATTACTACTTCCCACTTATTCCTATATATATATATATATTT	458
Db	1361	CGTAAGCTAGCGCTACGTTTCAAAGGCTCTTACCTACCTATCCTGTACAAGCTGTGCCGAAT	1420
Qy	459	TCAATATCTATTAATAGTAAAGCTTCATAGGGCTCTTCTGCTCCCTTAATATAAGAATCTCG	518
Db	1421	TCAATATCAGGCTACAGTAAAGCTCCAGGGGCTCTTCCGTCCTGTCGCGGGTAAACCTGC	1480
Qy	519	ATCTTCACAGATAATTTTATTTCAATTAAGATTTTTTTTAAAGACAGCATTTTAAAGTCGTTAC	578
Db	1481	ATCTTCACAGGTACTATGATTTCCCGAGTCTCTCGTTGAGACAGTGCCCAAAATCGTTAC	1540
Qy	579	ATCTTTTCATGCAAGTTC	594
Db	1541	GCCTTTTCGTCGGGTC	1556

RESULT 9

AAH54552	
ID	AAH54552 standard; DNA; 3030 BP.
XX	
XX	
AC	AAH54552;
XX	
DT	03-SEP-2001 (first entry)
XX	
DE	S. epidermidis genomlc polynucleotide sequence SEQ ID NO:3916.
XX	
KW	Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW	vaccination; endocarditis; ds.
XX	
OS	Staphylococcus epidermidis.
XX	
PN	WO200134809-A2.
XX	
PD	17-MAY-2001.
XX	
PF	09-NOV-2000; 2000WO-US30782.
XX	
PR	09-NOV-1999; 99US-0164258.
XX	
PA	(GLAX) GLAXO GROUP LTD.
XX	
PI	Kimmerly WJ;
XX	
DR	WPI; 2001-316495/33.

Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*, useful for vaccinating against infections, e.g. endocarditis -

Claim 8; Page 1571-1572; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH5090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH5091 to AAH5098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3030 BP; 918 A; 678 C; 519 G; 915 T; 0 other;

Query Match 32.0%; Score 190; DB 22; Length 3030;

Best Local Similarity 59.7%; Pred. No. 5.3e-20;

Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;

Qy	1	GTATCGCTTTAAATAGGCGGAACAGACTTACCCCTTAAACATACTACTGCTTTAGGATCGGA	60
Db	2225	GTACCGCTTTAATGGGCGAACAGCCCAACCCCTTGGGACCGGACTACAGCCCCAGGATCGGA	2284
Qy	61	TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGGGACTCTCGGAAAGATTTAGCC	120
Db	2285	TGAGCCGCACATCGAGGTGCCAAACCTTCCCGCTCGATGTGAACCTCTGGGGGAGATAAGCC	2344
Qy	121	TGTTATCCCTTAGAGTAACTTTTATCCGTTAAGCGATAATTTTATTATTAAATATATTATCG	180
Db	2345	TGTTATCCCGGGGTAGCTTTTATCCGTTAGCGATGGCCCTTCCATGTCCGGAACCAACCG	2403

DT	03-SEP-2001	(first entry)	
XX			
DE	S. epidermidis genomic polynucleotide sequence SEQ ID NO:3877.		
XX			
KW	Staphylococcus epidermidis SRI strain; infection; diagnosis;		
KW	vaccination; endocarditis; ds.		
XX			
OS	Staphylococcus epidermidis.		
XX			
PN	WO200134809-A2.		
XX			
PD	17-MAY-2001.		
XX			
PF	09-NOV-2000; 2000WO-US30782.		
XX			
PR	09-NOV-1999; 99US-0164258.		
XX			
PA	(GLAX) GLAXO GROUP LTD.		
XX			
PI	Kimmerly WJ;		
XX			
DR	WPI; 2001-316495/33.		
XX			
PT	Nucleic acids encoding polypeptides from Staphylococcus epidermidis,		
PT	useful for vaccinating against infections, e.g. endocarditis -		
XX			
PS	Claim 8; Page 1525-1526; 2188pp; English.		
XX			
CC	AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides		
CC	(II), given in AAG81434 to AAG83120, from Staphylococcus epidermidis.		
CC	(I) and (II) can have antibacterial activity and therefore can be used		
CC	in vaccination. The nucleic acids (I) may be used to produce the		
CC	S. epidermidis polypeptides (II) via the production of vectors		
CC	containing them which are used to produce hosts cells which express the		
CC	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be		
CC	used to vaccinate subjects and to raise antibodies against the bacteria.		
CC	The polypeptides may also be used to assay for other inhibitors of their		
CC	activity and therefore identify compounds that may be used for the		
CC	treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to		
CC	AAH55090 represent specifically claimed S. epidermidis genomic DNA		
CC	polynucleotide sequences from the present invention. AAH55091 to		
CC	AAH55098 represent oligonucleotide sequences and primers which are used		
CC	in the exemplification of the present invention.		
CC	N.B. The present invention specifically claims all the polynucleotide		
CC	sequences given in the sequence listing of the present specification,		
CC	however the sequence listing only goes up to SEQ ID NO:4454 so even		
CC	though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,		
CC	no sequences are present for SEQ ID NO:4455 to 4464.		
XX			
SQ	Sequence 3099 BP; 753 A; 860 C; 613 G; 873 T; 0 other;		
	Query Match 32.08; Score 190; DB 22; Length 3099;		
	Best Local Similarity 59.79; Pred. No. 5.3e-20;		
	Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;		
QY	1 GTATCGCTTTAATAGGGGACAGACTTACCCCTTAAACATACATACTACTGCTTAGGATGCGA 60		
Db			
1216	GTACCGCTTTAATGGCGACAGCCACCCCTTGGGACCGACTACAGCCCGAGGATGCGA 1275		
QY	61 TAAGCCGACATCGAGGTGCCAACCTTTTCGTCAATATGGACTCTCGGAAAGATATAGCC 120		
Db			
1276	TGAGCCGACATCGAGGTGCCAACCTCCCGTCGATGTGAACCTCTCTGGGGGAGATAAGCC 1335		
QY	121 TGTATCCCTAGACTACTTTTATTCGCTTAGCCGATATTTTATTAATAATTATCG 180		
Db			
1336	TGTATATCCCGGGTAGCTTTTATTCGCTTAGCCGATGCGCCTTCCCATGCGGAACACCG 1394		
QY	181 GATCATTAAGACCCGACATTAATCTCTGTTTAAATTTGTAATTTTACAGCTTAATATATAT 240		
Db			
1395	GATCAGCTAAGTCGCTTCTTCGACCCCTGCTCGACTTGAGGTCTCGCACTCAAGCTCCCTT 1454		
QY	241 TTATCTTTTATATATATATATAT-----AACATTGTACACCTCGC 278		
Db			

CC polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though the sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX
XX Sequence 3229 BP; 893 A; 611 C; 828 G; 897 T; 0 other;
SQ

Query Match 32.0%; Score 190; DB 22; Length 3229;
Best Local Similarity 59.7%; Pred. No. 5.3e-20;
Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;

QY 1 GTATCGCTTTAATAGGCGAACACAGACTTACCCTTAACATACTACTGCCTTAGAGATGGCA 60
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 2270 GTACCGCTTTAATGGCGAACAGCCAAACCTTGGGACCGACTACAGCCGCCAGGATGGGA 2211
QY 61 TAAGCGCATCGAGGTGCCAAAACCTTTTCGTCGAATATGGACTCTCGGAAAAAGATTAGCC 120
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 2210 TGAGCGGACATCGAGGTGCCAAAACCTCCCGTCGATGTGAACHTCTGGGGAGATAGAACC 2151
QY 121 TGTATCCCTAGAGTAACCTTTATCCGTTAAGCGATAATTTATTTAATAAATATATCG 180
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 2150 TGTATCCCGGGTAGCTTTATCCGTTGAGCGATGGCCT-TCCATCGGAACACCGC 2092
QY 181 GATCATTAAAGCCGACATTAATCTCTGTTTAATTTGTAATTTTACAGTTAATATATAT 240
DB ||||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 2091 GATCACTAAGTCGCGCTTTTGACCCCTCTCGACTTGTAGTCTCGCAGTCAAAGCTCCCTT 2032
QY 241 TTATCTTTTATATAATAATAT-----AACATTGTACACCTCCG 278
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 2031 ATGCGCTTTACACTCTATGAATGATTTCCAACCATCTGAGGGAACCTTTGAGCGCTCG 1972
QY 279 TTTTATATAGAGAGACGCCGCCAGCTAAACTATCTTATAAAATATGTTAAAAATTTT 338
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 1971 TTACCTTTTAGGCGGACGCCGCCAGCTAAACTGCGCGCTGACACTGTCTCCCACCCAC 1912
QY 339 GTTATAAAAAATTTTATAAGAAATTTATATATATAAAATGGTATTTTCATTAACAATTACA 398
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 1911 GATAAGTGGTGGGTTAGAAGCCACACAGCTAGGTAGTATCCACCACCGCTCCA 1852
QY 399 TTATTTCCAAAAAATAATATTAATACTACTTCCCAATTTCTATGTTATATATATATATT 458
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 1851 CGTAAGCTAGCGCTCACGTTTCAAGGCTCCTACCTATCCTGTACAAGCTGTGCGGAATT 1792
QY 459 TCAATATCTAATTAAGTAAAGCTTCATAGGCTTTCTGCTCTAATAATAAGAAATCTGC 518
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 1791 TCAATATCAGGCTACAGTAAAGCTCCACGGGGTCTTTCCGTCCTGTGCGGGGTAACTGC 1732
QY 519 ATCTTCACAGATAATTTTATTTCATTAAGATTTTTTTTTTAAGACAGCATTTAAGTCGTTAC 578
DB ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Dy 1731 ATCTTCACAGTACTATGATTTACCGAGTCTCTCGTTGAGACAGTGCCCAATCGTTAC 1672
QY 579 ATCTTTTCATGCGGTC 594
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 1671 GCCTTTTCGCGGGTC 1656

RESULT 13
AAH54853
ID AAH54853 standard; DNA; 3232 BP.
XX
XX AAH54853;
XX AC
XX DT
XX 03-SEP-2001 (first entry)
XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:4217.
XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX

Db 1388 GATAAGTGTGGGGTTAGAAAGCCAAACACAGCTAGGAGTAGTATCCCAACGCGCTCCA 1447
 QY 399 TTATTTCCAAATAAATAATATTACTTCCCATTTATTCCTATGTTATATATATATATTT 458
 Db 1448 CTAAGCTAGCGCTCAGCGTTTCAAGGCTTCCACCTATCTGTACAAAGCTGTGCGGAAT 1507
 QY 459 TCAATATCTATTAAATAGTAAAGCTTCATAGGCTCTTCTGTCTCTAATAAAGAAATCTGC 518
 Db 1508 TCAATATCAGGCTACAGTAAAGCTCCAGGGGTCTTTCGGTCTGTGCGGGGTAACTGTC 1567
 QY 519 ATCTTCACAGATAAATTTATTTTCATATAGATTTTTTTTAAAGACAGATTTAAGTGTGTTAC 578
 Db 1568 ATCTTCACAGGTACATATGATTTTCAACGAGTCTCTCGTTGAGACAGTGGCCAAATCGTTAC 1627
 QY 579 ATCTTTTCATGCGAGTTC 594
 Db 1628 GCCTTTCGTGCGGGTC 1643

RESULT 14
 AAH54539/c

ID AAH54539 standard; DNA: 3405 BP.

XX AAH54539;

AC AAH54539;

DT 03-SEP-2001 (first entry)

DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3903.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW Staphylococcus epidermidis; ds.

OS Staphylococcus epidermidis.

XX WO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

FA Kinmerly WJ;

PI WPI; 2001-316495/33.

DR Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

XX useful for vaccinating against infections, e.g. endocarditis -

PS Claim 8; Page 1556-1557; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,

CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 3405 BP; 971 A; 670 C; 922 G; 842 T; 0 other;

QY Query Match 32.0%; Score 190; DB 22; Length 3405;

Best Local Similarity 59.7%; Pred. No. 5.2e-20;

Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;

QY 1 GTATCGCTTTAATAGCGGAACAGACTTACCCCTTAACACATACACTGCTTGGATGCCA 60
 Db 2171 GTACCGCTTTAATGGCGAAGCAAGCCCAACCCCTTGGGACGACTACAGCCCCAGGATGCCA 2112
 QY 61 TAAGCGCATCGAGGTGCCAAACCTTTTCGTCATATGAGTCTCGGAAAAGATTAGCC 120
 Db 2111 TGAGCGCATCGAGGTGCCAAACCTCCCGTCGATGTGAACCTTTGGGGAGATAAGCC 2052
 QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATATATGTC 180
 Db 2051 TGTATCCCGGGGTAGCTTTTATCCGTTGAGCGATGGCCCT-TCCATCGGAACCCACCG 1993
 QY 181 GATCATTAAGACCGACATTAATCTCTGTGTTAATTTGTTAAATTTTACAGTTAATATATAT 240
 Db 1992 GATCATAAGTCCGCTTTTCGACCCCTGCTGACTGCTAGGTCTCGAGTCAAGCTCCCTT 1933
 QY 241 TTATCTTTATATAAATAATAT-----AACATTGTACACTCG 278
 Db 1932 ATGCGCTTTACACTCTATGAATGATTTCCAAACCATTTCTGAGGGAACCTTTGAGGCGCTCG 1873
 QY 279 TTTTATATAGGAGAGACCGCCCGCAGTCAAACTATCTTATAAATTTTAAAAATTTT 338
 Db 1872 TTACCTTTTAGGAGCGCGCCCGCAGTCAAACTGCCCCCTGACACTGTCTCCACCAC 1813
 QY 339 GTTATAAAAAATTTTATAAGAATTTATATATATAAATGATTTTCAATTACAAATTACA 398
 Db 1812 GATAAGTGTGCGGGTTAGAAAGCCCAACACAGCTAGGAGTAGTATCCCAACACCGCTCCA 1753
 QY 399 TTATTTTCCAAAAATAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 458
 Db 1752 CGTAAGCTAGCGCTCAGCGTTTCAAAAGGCTCCTACCTATCTGTACAGCTGTCCGAAT 1693
 QY 459 TCAATATCTATTAAATAGTAAAGCTTCATAGGCTCTTTCTGTCTTATATAAAGAAATCTGC 518
 Db 1692 TCAATATCAGGCTACAGTAAAGCTCCACGGGCTCTTTCGGTCTGTCTGCGGGTACCTGC 1633
 QY 519 ATCTTCACAGATAAATTTTATTTTCAATTAAGATTTTAAAGACAGCATTTAAGTCTGTTAC 578
 Db 1632 ATCTTCACAGTACTATGATTTTCAACGAGTCTCTCGTTGAGACAGTGCCTTACGTTAC 1573
 QY 579 ATCTTTTCATGCGAGTTC 594
 Db 1572 GCCTTTTCGTGCGGGTC 1557

RESULT 15

AAH54518/c

ID AAH54518 standard; DNA: 3500 BP.

XX AAH54518;

XX 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3882.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

XX Staphylococcus epidermidis.

XX WO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30782.

XX
PR 09-NOV-1999; 99US-0164258.

PA (GLAX)) GLAXO GROUP LTD.

PI Kimmerv WJ:

XX WPI: 2001-316495/33. DR

XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
PT useful for vaccinating against infections, e.g. endocarditis -

PS Claim 8; page 1530-1532; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AA8H1454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH53090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

SQ Sequence 3500 BP; 980 A; 712 C; 1004 G; 804 T; 0 other;

Query Match 32.0% Score 190; DB 22; Length 3500;

Query match: 52.0%; score: 190; DB 22;
Best Local Similarity 59.7%; Pred. No. 5.2e-20;

Best Local Similarity 39.7%, Freq. NO: 3.2E-20;
Matches 368; Conservative 0; Mismatches 225; Indels 23; Gaps 2;

QY 1 GTATCGCTTAA TAGGCGAACAGACTTACCCCTTAAACATAC TACTG CCTTAGGATGCCA 60

[illegible]

61 TAAGCCGACATCGAGGTGCCAACCTTTTCGTCAATATGGACTCTCGGAAAGATTAGCC 120

db 3323 TGAGCCGACATCGAGGTGCCAAACCTCCCGTCGATGTGAACCTCTTGGGGGAGATAAGCC 3264

121 TGTATCCCTAGAGTAACTTTTATCCGTAATAATTATTAATAATTATCG 180

db 3263 TCGTTATCCCCCGCGTACCTTTTATCCCTTTCAGCCATCCCCCT-TCCATCCCCGACCCACC 3205

[illegible][illegible][illegible][illegible]

0
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99

.....

11

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Q7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Q7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 10:11:59 : Search time 1849 seconds
(without alignments)
6722.749 Million cell updates/sec

Title: US-09-369-992c-1_copy_1147_1740

Perfect score: 594

Sequence: 1 gtagcttctaagcgaa.....ttacatcttctgacaggtc 594

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

PBU79731

LOCUS

DEFINITION

PBU79731

Plasmodium berghei extrachromosomal plasmid PB-1, ORF470 gene,

partial cds, tRNA-Thr, large subunit ribosomal RNA, tRNA-Met,

tRNA-Arg, tRNA-Val, tRNA-Arg, tRNA-Leu, tRNA-Ala, tRNA-Ala, and

small subunit ribosomal RNA genes, complete sequences.

U79731

U79731.1 GI:2662401

Plasmodium berghei.

Plasmodium berghei

Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

1 (bases 1 to 5849)

Yap,M.W., Kara,U.A., ten Heggeler-Bordier,B., Ting,R.C. and

Tan,T.M.

partial nucleotide sequence and organisation of extrachromosomal

plasmid-like DNA in Plasmodium berghei

Gene 200 (1-2), 91-98 (1997)

98038979

2 (bases 1 to 5849)

Yap,M.W.C., Kara,U.A.K. and Tan,T.M.C.

Direct Submission

1	594	100.0	5849	3	PBU79731	U79731 Plasmodium
2	578	97.3	594	3	AF182847	AF182847 Plasmodium
c	557.2	93.8	594	3	PVU97561	PVU97561 Plasmodium
4	554	93.3	594	3	AF182846	AF182846 Plasmodium
c	552.4	93.0	594	3	PVU97562	PVU97562 Plasmodium
c	546.8	92.1	593	3	PVU97565	PVU97565 Plasmodium
c	543	91.4	2700	3	PFLSRN	X61660 Plasmodium
c	543	91.4	5142	3	PFLSRN	X75545 P.falciparum
c	543	91.4	15421	3	PFCOMPRA	X95275 P.falciparum
c	541.4	91.1	595	3	PVU97560	U797560 Plasmodium
c	539.8	90.9	595	3	PVU97558	U797558 Plasmodium
c	532.6	89.7	594	3	PVU97559	U797559 Plasmodium
c	532	89.6	596	3	PVU97564	U797564 Plasmodium
c	530.4	89.3	596	3	PVU97563	U797563 Plasmodium
c	528.6	89.0	595	3	PVU97557	U797557 Plasmodium
c	512	86.2	598	3	PVU97556	U797556 Plasmodium
c	497	83.7	2621	3	PBU79732	U79732 Plasmodium
c	256.8	43.2	73345	8	AL0294725	U79732 Plasmodium
c	254	42.8	7291	1	AF211133	AF211133 Carsonell
c	252.6	42.5	34996	3	U87145	U87145 Toxoplasma
c	252.6	42.5	34996	3	U87145	U87145 Toxoplasma
c	249.4	42.0	2700	3	AF304316	AF304316 Neospora
c	249.4	42.0	7847	3	AF304319	AF304319 Neospora
c	249.2	42.0	6717	1	AF211135	AF211135 Carsonell
c	248.6	41.9	3269	3	AF304322	AF304322 Neospora
c	246.8	41.5	6730	1	AF211134	AF211134 Carsonell
c	242	40.7	5087	1	AF211143	AF211143 Carsonell
c	241.2	40.6	6113	3	TGTRNARRN	Y11430 T.gondii pl
c	240.8	40.5	3983	1	AF211144	AF211144 Carsonell
c	239.2	40.3	3987	1	AF211145	AF211145 Carsonell
c	236.2	39.8	164921	8	AF022186	AF022186 Cyanidium
c	235.6	39.7	7347	1	AF211124	AF211124 Carsonell
c	233.6	39.3	3985	1	AF280097	AF280097 Carsonell
c	233.6	39.3	3988	1	AF243137	AF243137 Carsonell
c	233.6	39.3	4569	1	AF211151	AF211151 Carsonell
c	232.4	39.1	7806	1	AF211141	AF211141 Carsonell
c	232	39.1	3986	1	AF211136	AF211136 Carsonell
c	232	39.1	3990	1	AF243138	AF243138 Carsonell
c	232	39.1	3996	1	AF211129	AF211129 Carsonell
c	232	39.1	4022	1	AF211148	AF211148 Carsonell
c	231.6	39.0	2957	8	CHPLIB	X61179 P.littorali
c	230.4	38.8	3987	1	AF243136	AF243136 Carsonell
c	230.4	38.8	3995	1	AF211125	AF211125 Carsonell
c	230.4	38.8	3998	1	AF211130	AF211130 Carsonell
c	228.8	38.5	3992	1	AF211131	AF211131 Carsonell

JOURNAL Submitted (26-NOV-1996) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore

FEATURES

source Location/Qualifiers

1..5849

/organism="Plasmodium berghei"

/strain="ANKA"

/db_xref="taxon:5821"

/note="extrachromosomal plasmid PB-1"

complement(<1..711)

/note="ORF470"

/codon_start=1

/protein_id="AAB88271.1"

/db_xref="GI:2662402"

/translation="MINKLKLKFLNLYNLYKQYNNKINLYLIRNGLNKLINLS
NNIYLPFFYKFIYSLKFLNWKFLDSCFPCNYSYNNIYYSTLKDLSNLLYLR
TNLNIINFYFNKRIIPIDVCDISVILHTQYFLKLGILIFLSLLDVLKPLVLR
KYLGVSYVDNPFANINISIISEGSFCYVSKYIKGSPNLSYFKTNSYDFAQFERTL
LIASEFAYGVLEGCTASL"

complement(739..809)

/product="trna-Thr"

/product="trna-Thr"

complement(814..3521)

/product="large subunit ribosomal RNA"

complement(3565..3636)

/product="trna-Met"

complement(3646..3717)

/product="trna-Arg"

3742..3813

/product="trna-Val"

3817..3888

/product="trna-Arg"

3895..3975

/product="trna-Leu"

complement(3996..4058)

/product="trna-Asn"

4089..4159

/product="trna-Ala"

4208..5648

/product="small subunit ribosomal RNA"

5793..5849

/product="trna-Ile"

BASE COUNT 2296 a 673 c 557 g 2323 t

ORIGIN

Query Match 100.0%; Score 594; DB 3; Length 5849;
Best Local Similarity 100.0%; Pred. No. 4,1e-73;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACATACTACTGCGCTTAGGATGCGA 60
Db 1147 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACATACTACTGCGCTTAGGATGCGA 1206

Qy 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCGAATATGACTCTCGGAAAGATTAGCC 120
Db 1207 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCGAATATGACTCTCGGAAAGATTAGCC 1266

Qy 121 TGTATCCCTAGAGTAACCTTTATCCGCTTAAGCGATAATTTTATTATTAATAAATTATCG 180
Db 1267 TGTATCCCTAGAGTAACCTTTATCCGCTTAAGCGATAATTTTATTATTAATAAATTATCG 1326

Qy 181 GATCATTAAAGACCGACATTAATCTCTGTTTAAATTTGTTAAATTTTACAGTTAATATATAT 240
Db 1327 GATCATTAAAGACCGACATTAATCTCTGTTTAAATTTGTTAAATTTTACAGTTAATATATAT 1386

Qy 241 TTATCTTTATATATAATAATAACATGTACACCTCGGTTTTTATATAGGAGGAGACCGC 300
Db 1387 TTATCTTTATATATAATAATAACATGTACACCTCGGTTTTTATATAGGAGGAGACCGC 1446

Qy 301 CCCAGTCAAACTACTCTTATAAATTTGTTAAATTTTGTATAAATAATTTTATAAGAAT 360
Db 1447 CCCAGTCAAACTACTCTTATAAATTTGTTAAATTTTGTATAAATAATTTTATAAGAAT 1506

Qy 361 TTATATATATATAAATGGTATTTTCATTAACAATTAATTTCCAAAAAATAATATT 420

Db 1507 TTATATATATAAATGGTATTTTCATTAACAATTAATTTCCAAAAAATAATATT 1566

Qy 421 ACTACTTCCCATTTATCTATGTTATATATATATATATATATTTCAATATCTATTAATAGTAAAG 480

Db 1567 ACTACTTCCCATTTATCTATGTTATATATATATATATATTTTCAATATCTATTAATAGTAAAG 1626

Qy 481 CTTTCATAGGCTCTTCTGCTCTAATAAAGAAATCTGCATCTTCACAGATAATTTTATTT 540

Db 1627 CTTTCATAGGCTCTTCTGCTCTAATAAAGAAATCTGCATCTTCACAGATAATTTTATTT 1686

Qy 541 CATTAAAGATTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCGAGTC 594

Db 1687 CATTAAAGATTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCGAGTC 1740

RESULT 2

AF182847

LOCUS

DEFINITION

Plasmodium chabaudi from Australia large subunit ribosomal RNA

AF182847

gene, partial sequence.

ACCESSION

AF182847

VERSION

AF182847.1 GI:6110458

SOURCE

Plasmodium chabaudi.

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE

1 (bases 1 to 594)

Tham, J.M., Khoo, L.K. and Kara, A.U.K.

AUTHORS

Direct Submission

TITLE

Submitted (03-SEP-1999) Institute of Molecular and Cell Biology, 30

JOURNAL

Medical Drive, Singapore 117609, Singapore

FEATURES

Location/Qualifiers

source

1..594

/organism="Plasmodium chabaudi"

/organelle="plastid"

/strain="adami DS"

/db_xref="taxon:5825"

/country="Australia"

<1..>594

/product="large subunit ribosomal RNA"

BASE COUNT 205 a 95 c 69 g 225 t

ORIGIN

Query Match 97.3%; Score 578; DB 3; Length 594;
Best Local Similarity 98.3%; Pred. No. 1e-70;
Matches 584; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACATACTACTGCGCTTAGGATGCGA 60
Db 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACATACTACTGCTTTAGGATGCGA 60

Qy 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCGAATATGACTCTCGGAAAGATTAGCC 120
Db 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCGAATATGACTCTCGGAAAGATTAGCC 120

Qy 121 TGTATCCCTAGAGTAACCTTTATCCGCTTAAGCGATAATTTTATTATTAATAAATTATCG 180
Db 121 TGTATCCCTAGAGTAACCTTTATCCGCTTAAGCGATAATCTTATTATTAATAAATTATCG 180

Qy 181 GATCATTAAAGACCGACATTAATCTCTGTTTAAATTTGTTAAATTTTACAGTTAATATATAT 240
Db 181 GATCATTAAAGACCGACATTAATCTCTGTTTAAATTTGTTAAATTTTACAGTTAATATATAT 240

Qy 241 TTATCTTTATATATAATAATAACATGTACACCTCGGTTTTTATATAGGAGGAGACCGC 300
Db 241 TTATCTTTATATATAATAATAACATGTACACCTCGGTTTTTATATAGGAGGAGACCGC 300

Qy 301 CCCAGTCAAACTACTCTTATAAATTTGTTAAATTTTGTATAAATAATTTTATAAGAAT 360
Db 301 CCCAGTCAAACTACTCTTATAAATTTGTTAAATTTTGTATAAATAATTTTATAAGAAT 360

Qy 361 TTATATATATATAAATGGTATTTTCATTAACAATTAATTTCCAAAAAATAATATT 420

Db	354	ATATCTTTATATACATATATAA	CTTGTGTA	CTCTCTCGTTTTTATATATAGGAGGAGACCGC	295
Qy	301	CCCACTCAAACTATCTTTAT	AAATATCTGTTAAAAATTTT	GTATATAAAAAATTTTATATAAGAAAT	360
Db	294	CCCACTCAAACTATCTTTAT	AAATAC	TCTTAAAAATTTTGTAAAAAATTTTTTTTAAGAAT	235
Qy	361	TTATATATATATAAAATGGT	ATATTTTCATTACAAAT	TACATATTTTCCAAAAAATAATATT	420
Db	234	TTATATATATAAAAAATGGT	ATATTTTCATTACAAAT	TACATATTTTCCAAAAAATAATATT	175
Qy	421	ACTACTTCCCATTTATTTCT	TATGTTATATATATATAT	TTTTCAAATATCTTATTAATAGTAAAG	480
Db	174	ATTATTTCCCATTTATGCT	ATGCTATATATATATATAT	TTTTTCAATATCTTATTCATAGTAAAG	115
Qy	481	CTTCATAGGGTCTTTCTG	TCTCTATATATAAAGAAATCT	GCATCTTCCACAGATAATTTTATT	540
Db	114	CTTCATAGGGTCTTTCTG	TCTCTATATATAAAGAAATCT	GCATCTTCCACAGATAATTTTATT	55
Qy	541	CATTAAAGATTTTTTTTAA	GACAGCATTTAAGTCTGTACATCT	TTTTCATGCGAGGTC	594
Db	54	CATTAAAGATTTTTTTTAA	GACAGCATTTAAGTCTGTACATCT	TTTTCATGCGAGGTC	1
RESULT 4					
AF182846					
LOCUS					
DEFINITION					
AF182846 Plasmodium malariae from Burma large subunit ribosomal RNA gen					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
1..594					
/organism="Plasmodium malariae"					
/organelle="plastid"					
/strain="BTTSH/104"					
/db_xref="taxon:5858"					
/country="Burma"					
<1..>594					
/product="large subunit ribosomal RNA"					
BASE COUNT	201	a	90	c	70
ORIGIN			g	233	t
Query Match 93.3%; Score 554; DB 3; Length 594;					
Best Local Similarity 95.8%; Pred. No. 2e-67;					
Matches 569; Conservative 0; Mismatches 25; Indels 0; Gaps					
Qy	1	GTATCGCTTTTAATAGGCGA	ACAGACTTACCCCTTAAAA	CATACACTACTGCTTAGGATCGCA	60
Db	1	GTATCGCTTTTAATAGGCGA	ACAGACTTACCCCTTAAAA	CATACACTACTGCTTAGGATGCGA	60
Qy	61	TAAAGCCGACATCGAGGTG	CGCAACCTTTTTCGTC	CAATATGACCTCTCGGAAAAAGATTAGCC	120
Db	61	TAAAGCCGACATCGAGGTG	CGCAACCTTTTTCGTC	CAATATGACCTCTCGGAAAAAGATTAGCC	120
Qy	121	TGTTATCCCTAGAGTAACT	TTTTTATCCGTTAAGCGTA	AAATTTTATTTAAATAATATTCG	180
Db	121	TGTTATCCCTAGAGTAACT	TTTTTATCCGTTAAGCGTA	AAATTTTATTTAAATAATATTCG	180
Qy	181	GATCATTAAAGACCGCAG	CATTAAATCTCTGTTTAA	TTTTCGTAATTTTACAGTTAATATATAT	240
Db	181	GATCATTAAAGACCGCAG	CATTAAATCTCTGTTTAA	TTTTCGTAATTTTACAGTTAATATATAT	240
Qy	241	TTATCTTTATATAATAA	TATAACATTTGTACACCT	CCCGTTTTTATATAGGAGGAGACCGC	300
Db	241	TTATCTTTATATAATAA	TATAACATTTGTACACCT	CCCGTTTTTATATAGGAGGAGACCGC	300

Db 241 ATATCTTTTATAATAAATAAATTAACATTTGTAATCTCTCGTTTATATATATAGGAGGAGCCGC 300
QY 301 CCACGCTCAAACTATCTTATAAATATGTTAAATAATTTGTTATAAAAAATTTTATAAGAAT 360
Db 301 CCACGCTCAAACTATCTTATAAATATGTTAAATAATTTGTTATAAAAAATTTTATAAGAAT 360
QY 361 TTATATATATAATAAATGTTATTTCAATTAACAATTAACATTTTATTTCCAAAAAATAATATT 420
Db 361 TTATATATATAAAAAATGTTATTTCAATTAACAATTTTATTTCCAAAAAATAATATT 420
QY 421 ACTACTTCCCATTTATTTCTATGTTATATATATATATTTTCAATATCTATTAATAGTAAAG 480
Db 421 TATATTTCCCATTTATTTCTATGTTATATATATATTTTCAATATCTATTAATAGTAAAG 480
QY 481 CTTATAGGGCTTTCTGCTCTAATAAAGAAATCTGCATCTTTCACAGATAAATTTTATT 540
Db 481 CTTATAGGGCTTTCTGCTCCCAATATAAAGAAATCTGCATCTTTCACAGATAAATTTTATT 540
541 CATTAAGAATTTTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCAGGTC 594
Db 541 CATTAAGAATTTTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCAGGTC 594

RESULT 5
PVU97562/c
LOCUS
DEFINITION
Plasmodium vivax extrachromosomal plastid large subunit ribosomal
RNA gene, partial sequence.
ACCESSION
U97562
VERSION
U97562.1 GI:4100400
KEYWORDS
malaria parasite P. vivax.
ORGANISM
Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 594)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct PCR amplification and sequence analysis of extrachromosomal
Plasmodium DNA from dried blood spots
JOURNAL
Acta Trop. 68 (1), 105-114 (1997)
MEDLINE
98013247
PUBMED
9352006
REFERENCE
2 (bases 1 to 594)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct Submission
JOURNAL
Submitted (15-APR-1997) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore
FEATURES
Location/Qualifiers
1. .594
/organism="Plasmodium vivax"
/isolate="primary isolate pv13/p"
/db_xref="taxon:5855"
/note="extrachromosomal plastid"
<1..>594
/product="large subunit ribosomal RNA"
BASE COUNT 225 a 72 c 99 g 198 t
ORIGIN

Query Match 93.0%; Score 552.4; DB 3; Length 594;
Best Local Similarity 95.8%; Pred. No. 3.4e-67;
Matches 568; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATCTACTCTGCTTAGGATGCGA 60
Db 594 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATCTACTCTGCTTAGGATGCGA 535
QY 61 TAAGCCGACATCGAGTGCAGAACCTTTTCGTCGAATATGACHTCTCGGAAAAGATTAGCC 120
Db 534 TAAGCCGACATCGAGTGCAGAACCTTTTCGTCGAATATGACHTCTCGGAAAAGATTAGCC 475
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
Db 474 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 415

QY 181 GATCATTAAAGACCGACATTAATCTCTGTTAAATTTGTAATAATTTTACAGTTAAATATATAT 240
Db 414 GATCATTAAAGACCGACATTAATCTCTGTTAAATTTGTAATAATTTTACAGTTAAATATAT 355
QY 241 TTATCTTTTATATAATAAATAACATTTGTACACCTCGTTTTTATATAGGAGGAGACCGC 300
Db 354 ATATCTTTTATATAAATATTAATCTCTGTTAAATTTGTAATAATTTTACAGTTAAATATAT 295
QY 301 CCACGCTCAAACTATCTTATAAATATGTTAAATAATTTGTTATAAAAAATTTTATAAGAAT 360
Db 294 CCACGCTCAAACTATCTTATAAATATGTTAAATAATTTGTTATAAAAAATTTTATAAGAAT 235
QY 361 TTATATATATAATAAATGTTATTTCAATTAACAATTAACATTTTCCAAAAAATAATATT 420
Db 234 TTATATATATAAAAAATGTTATTTCAATTAACAATTAACATTTTCCAAAAAATAATATC 175
QY 421 ACTACTTCCCATTTATTTCTATGTTATATATATATTTTCAATATCTATTAATAGTAAAG 480
Db 174 ATTATTTCCCATTTATTTCTATGTTATATATATATTTTCAATATCTATTAATAGTAAAG 115
QY 481 CTTATAGGGCTTTCTGCTCTAATAAAGAAATCTGCATCTTTCACAGATAAATTTTATT 540
Db 114 CTTATAGGGCTTTCTGCTCTAATAAAGAAATCTGCATCTTTCACAGATAAATTTTATT 55
QY 541 CATTAAGAATTTTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCAGGTC 594
Db 54 CATTAAGAATTTTTTTTAAAGACAGCATTTAAAGTCGTTACATCTTTTCATGCAGGTC 1

RESULT 6
PMU97565/c
LOCUS
DEFINITION
Plasmodium malariae extrachromosomal plastid large subunit
ribosomal RNA gene, partial sequence.
ACCESSION
U97565
VERSION
U97565.1 GI:4100403
KEYWORDS
Plasmodium malariae.
ORGANISM
Plasmodium malariae
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 593)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct PCR amplification and sequence analysis of extrachromosomal
Plasmodium DNA from dried blood spots
JOURNAL
Acta Trop. 68 (1), 105-114 (1997)
MEDLINE
98013247
PUBMED
9352006
REFERENCE
2 (bases 1 to 593)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct Submission
JOURNAL
Submitted (15-APR-1997) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore
FEATURES
Location/Qualifiers
1. .593
/organism="Plasmodium malariae"
/isolate="primary isolate pv16/1"
/db_xref="taxon:5858"
/note="extrachromosomal plastid"
<1..>593
/product="large subunit ribosomal RNA"
BASE COUNT 227 a 70 c 98 g 198 t
ORIGIN

Query Match 92.1%; Score 546.8; DB 3; Length 593;
Best Local Similarity 96.1%; Pred. No. 2e-66;
Matches 571; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATCTACTCTGCTTAGGATGCGA 60
Db 593 GTATCGCTTTAATAGGCGAACAGACTTACCCCTTAAACATCTACTCTGCTTAGGATGCGA 534

QY 61 TAAGCCGACATCGAGGTGCGCAACCTTTTCGTCATATGAGACTCTCGGAAAGATTAGCC 120
|||||
Db 533 TAAGCCGACATCGAGGTGCGCAACCTTTTCGTCATATGAGACTCTCGGAAAGATTAGCC 474
QY 121 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTATCG 180
|||||
Db 473 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTATCG 414
QY 181 GATCATTAAAGACCGACATTAACTCTGTTTAAATTTTACAGTTAAATATATAT 240
|||||
Db 413 GATCATTAAAGACCGACATTAACTCTGTTTAAATTTTACAGTTAAATATATAT 354
QY 241 TTATCTTTTATTAATAATATACATGTCACCTCGGTTTATATAGAGGAGACCGC 300
|||||
Db 353 ATATCTTTTATTAATAATATACATGTCACCTCGGTTTATATAGAGGAGACCGC 294
QY 301 CCAGTCAACATCTATTAATAATATGTTAAATAATTTTGTATATAAATTTTATAAGAT 360
|||||
Db 293 CCAGTCAACATCTATTAATAATATGTTAAATAATTTTGTATATAAATTTTATAAGAT 235
QY 361 TTATATATATAAATAATGTTTATTAACAAATTTACATTTTCCAAAAAATAATAT 420
|||||
Db 234 TTATATATATAAATAATGTTTATTAACAAATTTACATTTTCCAAAAAATAATAT 175
QY 421 ACTACTTCCATTTATCTATGTTATATATATATTTTCAATATCTATATAGTAAG 480
|||||
Db 174 ATTATTTCCATTTATGCTATGCTCTATATATATATTTTCAATATCTATATAGTAAG 115
QY 481 CTTATAGGCTCTTCTGTCCTAATAATAAGAAATCGCATCTTCACAGATAATTTAT 540
|||||
Db 114 CTTATAGGCTCTTCTGTCCTAATAATAAGAAATCGCATCTTCACAGATAATTTAT 55
QY 541 CATTAAAGATTTTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTC 594
|||||
Db 54 CATTAAAGATTTTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTC 1

RESULT 7
PFLLSRN/c
LOCUS
DEFINITION
P. falciparum gene for a large subunit ribosomal RNA from
the inverted repeat within the 35-Kb circular DNA.
X61660
VERSION
X61660.1 GI:13318
KEYWORDS
inverted repeat; ribosomal RNA; ribosomal RNA large subunit.
SOURCE
malaria parasite P. falciparum.
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
REFERENCE
1 (bases 1 to 2700)
Gardner, M.J.
Direct Submission
Submitted (22-AUG-1991) M.J. Gardner, Div of Parasitology, National
Inst for Medical Research, The Ridgeway - Mill Hill, London NW7
1AA, UK
2 (bases 1 to 2700)
Gardner, M.J., Feagin, J.E., Moore, D.J., Rangachari, K.,
Williamson, D.H. and Wilson, R.J.
Sequence and organization of large subunit rRNA genes from the
extrachromosomal 35 kb circular DNA of the malaria parasite
Plasmodium falciparum
Nucleic Acids Res. 21 (5), 1067-1071 (1993)
93219063
JOURNAL
MEDLINE
FEATURES
source
Location/Qualifiers
1. .2700
/organism="Plasmodium falciparum"
/strain="C10"
/db_xref="taxon:5833"
/dev_stage="Erythrocytic"
/clone="pf4, pf5, pf6a"
1. .2700
/note="Subcellular localisation unknown"
/product="large subunit rRNA"
1175 a 199 c 324 g 1002 t
rRNA
BASE COUNT

ORIGIN

Query Match 91.4%; Score 543; DB 3; Length 2700;
Best Local Similarity 95.6%; Pred. No. 4.8e-06;
Matches 569; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

QY 1 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 60
|||||
Db 2376 GTATCGCTTTAATAGCGAACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 2317
QY 61 TAAGCCGACATCGAGGTGCGCAACCTTTTCGTCATATGAGACTCTCGGAAAGATTAGCC 120
|||||
Db 2316 TAAGCCGACATCGAGGTGCGCAACCTTTTCGTCATATGAGACTCTCGGAAAGATTAGCC 2257
QY 121 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTATCG 180
|||||
Db 2256 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTATCG 2197
QY 181 GATCATTAAAGACCGACATTAACTCTGTTTAAATTTTACAGTTAAATATATAT 240
|||||
Db 2196 GATCATTAAAGACCGACATTAACTCTGTTTAAATTTTACAGTTAAATATATAT 2137
QY 241 T-TATCTTTTATTAATAATATACATGTCACCTCGGTTTATATAGAGGAGACCG 299
|||||
Db 2136 TATATCTTTTATTAATAATATACATGTCACCTCGGTTTATATAGAGGAGACCG 2077
QY 300 CCCAGTCAACATCTATTAATAATGTTAAATAATTTTGTATATAAATTTTATAAGAA 359
|||||
Db 2076 CCCAGTCAACATCTATTAATAATGTTAAATAATTTTGTATATAAATTTTATAAGAA 2017
QY 360 TTATATATATAAATAATGTTTAAACAAATTTACATTTTCCAAAAAATAATAT 419
|||||
Db 2016 TTATATATAAATAATGTTTAAACAAATTTACATTTTCCAAAAAATAATAT 1957
QY 420 TACTACTTCCATTTATCTATGTTATATATATATATTTTCAATATCTATTAATAGTAA 479
|||||
Db 1956 TATGTGTTCCCATTTATCTATGTTAAATATATATATTTTCAATATTTTATTAAGTAA 1897
QY 480 GCTTCATAGGCTCTTCTGTCCTAATAATAAGAAATCGCATCTTCACAGATAATTTAT 539
|||||
Db 1896 GCTTCATAGGCTCTTCTGTCCTAATAATAAGAAATCGCATCTTCACAGATAATTTAT 1837
QY 540 TCATTAAAGATTTTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTC 594
|||||
Db 1836 TCATTAAAGATTTTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTC 1782

RESULT 8

PFTRNA
LOCUS
DEFINITION
P. falciparum gene for tRNA I, A, N, L, R, V, R, M, T.
X75545
ACCESSION
X75545.1 GI:520908
VERSION
KEYWORDS
large ribosomal subunit; small ribosomal subunit RNA; transfer
RNA-Ala; transfer RNA-Arg; transfer RNA-Asn; transfer RNA-Ile;
transfer RNA-Leu; transfer RNA-Thr; transfer RNA-Val; tRNA gene;
trnI gene; trnL gene; trnM gene; trnN gene; trnR gene; trnS gene;
trnV gene.
SOURCE
malaria parasite P. falciparum.
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
REFERENCE
1 (bases 1 to 5142)
Gardner, M., Preiser, P., Rangachari, K., Moore, D., Feagin, D.,
Williamson, D.H. and Wilson, R.J.
Nine duplicated tRNA genes on the plastid-like DNA of the malaria
parasite Plasmodium falciparum
Gene 140, 307-308 (1994)
REFERENCE
2 (bases 1 to 5142)
Wilson, R.J.M.
Direct Submission
Submitted (02-NOV-1993) R.J.M. Wilson, National Institute for
Medical Research, Mill Hill, London NW7 1AA, UK
LOCATION/Qualifiers
FEATURES

source 1. .5142
/organism="Plasmodium falciparum"
/strain="BW (C10)"
/db_xref="taxon:5833"
/chromosome="35kb circular DNA"
/dev_stage="erythrocytic"
/clone="HindIII (frag5)"
complement(16. .89)
/gene="trnI"
/product="transfer RNA-Ile"
complement(16. .89)
/gene="trnI"
complement(237. .1663)
/gene="SSU rRNA"
/product="small subunit ribosomal RNA"
complement(237. .1663)
/gene="SSU rRNA"
complement(1710. .1780)
/gene="trnA"
/product="transfer RNA-Ala"
complement(1710. .1780)
/gene="trnA"
1810. .1881
/gene="trnN"
/product="transfer RNA-Asn"
1810. .1881
/gene="trnN"
complement(1891. .1970)
/gene="trnL"
/product="transfer RNA-Leu"
complement(1891. .1970)
/gene="trnL"
complement(1978. .2049)
/gene="trnR"
/product="transfer RNA-Arg"
complement(1978. .2049)
/gene="trnR"
complement(2051. .2122)
/gene="trnV"
/product="transfer RNA-Val"
complement(2051. .2122)
/gene="trnV"
2142. .2214
/gene="trnR"
/product="transfer RNA-Arg"
2142. .2214
/gene="trnR"
2220. .2293
/gene="trnM"
/product="transfer RNA-Met"
2220. .2293
/gene="trnM"
2335. .5030
/gene="LSU rRNA"
/product="large subunit ribosomal RNA"
2335. .5030
/gene="LSU rRNA"
5045. .5117
/gene="trnT"
/product="transfer RNA-Thr"
5045. .5117
/gene="trnT"
BASE COUNT 2078 a 495 c 561 g 2007 t 1 others
ORIGIN

Query Match 91.4%; Score 543; DB 3; Length 5142;
Best Local Similarity 95.6%; Pred. No. 4.2e-66;
Matches 569; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
QY 1 GTATCGCTTTAATAGCGAAGACAGACTTACCTTAAACATACACTACCTAGGATGCGA 60
|||||
DB 4705 GTATCGCTTTAATAGCGAAGACAGACTTACCTTAAACATACACTACCTTAGATGCGA 4646
|||||

QY 61 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGAGACTCTCGGAAAGATTAGCC 120
Db 4645 TAAGCCGACATCGAGGTGCCAAACCTTTTCGTCAATATGAGACTCTCGGAAAGATTAGCC 4586
QY 121 TGTATCCCTAGAGTAACTTTTATCCGTTAAGCGCAATTTTATATTAATAATTATCG 180
Db 4585 TGTATCCCTAGAGTAACTTTTATCCGTTAAGCGCAATTTTATATTAATAATTATCG 4526
QY 181 GATCATTAAAGACCGACATTAATCTCTCTTTAAATTTGTAATAATTTTACAGTAAATATAT 240
Db 4525 GATCATTAAAGACCGACATTAATCTCTCTTTAAATTTGTAATAATTTTACAGTAAATATAT 4466
QY 241 T-TATCTTTATATAATAATAATAACATTTGACACCTCCGTTTATATATAGAGGAGACCG 299
Db 4465 TATATCTTTATATAATAATAATAACATTTGACACCTCCGTTTATATATAGAGGAGACCG 4406
QY 300 CCCAGTCAAACTATCTTATAAATAATTTGTAATAATTTTGTATATAAATAATTTTATAAGAA 359
Db 4405 CCCAGTCAAACTATCTTATAAATAATTTGTAATAATTTTGTATATAAATAATTTTATAAGAA 4346
QY 360 TTTATATATATAAATAATGTTTTCATTAACAATTTACATTTTCCAAAAATAATAT 419
Db 4345 TTTATATATAAATAATGTTTTCATTTTAACTAAATTTTCCAAAGAAATAATAT 4286
QY 420 TACTACTTCCATTTATCTTATATATATATATATATATTTCAATATCTTATTAATAGTAAA 479
Db 4285 TATTGTTTCCCATTTATCTATGTTTAAATATATATATATTTTCAATATTTTATTAATAGTAAA 4226
QY 480 GCTTCATAGGCTCTTCTGTCCTAAATATAGAAATCGCATCTTCACAGATAATTTTATT 539
Db 4225 GCTTCATAGGCTCTTCTGTCCTAAATATAGAAATCGCATCTTCACAGATAATTTTATT 4166
QY 540 TCATTAAAGATTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTTC 594
Db 4165 TCATTAAAGATTTTAAAGACAGCATTTAAGTCGTTTACATCTTTTCATGCGAGTTC 4111

RESULT 9
PFCONSPIRA/c PFCONSPIRA 15421 bp DNA linear INV 14-FEB-1997
LOCUS P. falciparum complete gene map of plastid-like DNA (IR-A).
DEFINITION X95275
ACCESSION X95275
VERSION X95275.1 GI:1171583
KEYWORDS LSU rRNA gene; ORF 101; ORF470; ORF51; rpoB gene; rpoC gene; rpoD gene; rps2 gene; SSU rRNA gene; trnA-Ala; trnA-Arg; trnA-Asn; trnA-Ile; trnA-Leu; trnA-Met; trnA-Val.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
REFERENCE 1 (bases 1 to 15421)
AUTHORS Wilton, R.J.M., Denny, P.W., Preiser, P.R., Rangachari, K., Roberts, K., Roy, A., Whyte, A., Strath, M., Moore, D.J., Moore, P.W. and Williamson, D.H.
TITLE Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium falciparum
J. Mol. Biol. 261 (2), 155-172 (1996)
MEDLINE 96346169
REFERENCE 2 (bases 1 to 15421)
AUTHORS Wilton, R.J.M.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1996) R.J.M. Wilton, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK
COMMENT Related sequences X52177, X57167, X61660, X74308, X75544, and X75545.
FEATURES Location/Qualifiers
1..15421
/organism="Plasmodium falciparum"
/strain="C10"
/db_xref="taxon:5833"
/dev_stage="erythrocytic"
/note="IR-A half of 35kb circle, putative Plastid DNA"
complement(17..74)
/gene="trnA-Ile"
trnA

```
gene /product="tRNA-Ile"
      complement(17..74)
rRNA /gene="tRNA-Ile"
      complement(236..1662)
gene /product="small subunit ribosomal RNA"
      complement(236..1662)
tRNA /gene="SSU rRNA"
      complement(1709..1779)
gene /product="tRNA-Ala"
      complement(1709..1779)
tRNA /gene="tRNA-Ala"
      complement(1809..1871)
gene /product="tRNA-Asn"
      complement(1891..1969)
tRNA /product="tRNA-Asn"
      complement(1891..1969)
gene /product="tRNA-Leu"
      complement(1891..1969)
tRNA /product="tRNA-Leu"
      complement(1977..2048)
gene /product="tRNA-Arg"
      complement(1977..2048)
tRNA /product="tRNA-Arg"
      complement(2050..2121)
gene /product="tRNA-Val"
      complement(2050..2121)
tRNA /product="tRNA-Val"
      complement(2142..2213)
gene /product="tRNA-Arg"
      complement(2142..2213)
tRNA /product="tRNA-Arg"
      complement(2219..2292)
gene /product="tRNA-Met"
      complement(2219..2292)
rRNA /product="large subunit ribosomal RNA"
      complement(2335..5116)
gene /product="LSU rRNA"
      complement(2335..5116)
gene /product="ORF470"
      complement(5142..6554)
CDS /gene="ORF470"
      complement(5142..6554)
      note="red algal chloroplast"
      /codon_start=1
      /protein_id="CAA64569.1"
      /db_xref="GI:1171584"
      /db_xref="SPTREMBL:Q25799"
      /translation="MIKLNFLNIYLNKYQYKKNLNLVLIROGLNLINLKNLSSNI
      FLYMFIYFKKYSKLLNIPKLPDWNFDPCNINYDNIYYSSILKDNLLIYLYKNNL
      NIEFLSDKNFANISIIIFSGSFYIPKYKCNFLSTYFKTSSDFAQFERTLIIVG
      KYSVYEGCTASLYKESQLHVAIVEIIVKDYIKYTYLQNWYRGDYLNGGLYNF
      TTRGICLNTSKLDIOVEGSIITWKYPSTILKGFESISNFYSISNMQIADTGS
      KMWHIGSYTKSYIISKISLNSNLNIFGLVILKPFYSKYNKYTECSLLIFGNSLTV
      TIPYKYNWTSYKQAFVSKTEIIVLFLMQRGLSISSESISLIIIGFCSDIYNKLP
      FERNLEIPILFSLKIDIFN"
      /protein_id="6561..6716"
      /gene="ORF51"
      /db_xref="ORF51"
      /codon_start=1
      /protein_id="CAA64570.1"
      /db_xref="GI:1171585"
```

```
gene
CDS
```

```
/db_xref="SPTREMBL:Q25800"
/translation="MKYNKNVINTNFLVINKYKKLLKYKKNKNIYIILYKKYVILILKK
LIVLNI"
6725..7030
/gene="ORF101"
6725..7030
/gene="ORF101"
/codon_start=1
/protein_id="CAA64571.1"
/db_xref="GI:1171586"
/db_xref="SPTREMBL:Q25801"
/translation="MIKNIKKYKYKKNFKQKIIINKKFIINKNLIILKLPFYEY
IKNKLQIFNFKGILLNKNNNYIILLAFNNNDIIVLYFNLTNFINIKFKGKNIN"
7031..10105
/gene="rpob"
7031..10105
/gene="rpob"
/codon_start=1
/protein_id="CAA64572.1"
/db_xref="GI:1171587"
/db_xref="SWISS-PROT:P21421"
/translation="MIYIVNPIVKNYIISNLYLLLIQELIYNLYRYIYLFNNNINW
KFNFIYKIIILLNINISDITQINNNLKIILTLKLNINKKIIFKILFILP
FIYNNIILNGLYKTCQLEKKNKFIIFKFNKNKIYIYVYISGLRIIFKISKL
NIDCFNNKFNFLIILLYNNIYINKNISLIYNNIINKKIILYNTFYKYSYNNI
NNIISLKIIFKLNKNNIYNLLNLIIFSKLNFSYSDFYINNINPKFYSIDNNLL
IKSKYLKPKYQLLNINNIYNNITLLNKKYNIITILENININPLVOYSDOVNLL
EINQFKNMITTGLNSKFIILNDLRELPRNLIYISLINTNEGTLGVLNLTNIF
LNKILFYIYKHYFNPNLFLNFKNFYNSFNNIYKKNINFKNTIITLITNIN
NFKICNTQNIYIPFNLSFIENLPIFYHYNDSIRNLSIKKHQIVPIYFNLS
NIITYNFILNKLHIIISYQEGIVYVSCIKIIRDLFNQIYLYNNYKKIINONI
LIYKPIVWGEKVNIGIILAINNLLNSVSLGNLLVGYSLVGYEYDAIISRK
ILYNNLTSLNIYEISLNIINNIPEICSINLSKMYKNIKHLDKYILKEGYILA
NNILSKLMPFPFNKSLININFLPGSKLRFKPKPIISTHIDIGRVKIEILPN
HLYNTEKNIIYLAIRIYIQKYLQGLDKICNRHGHKGIISYISEINDPILNKKIQ
PDIFISAISPSRINIQIFEGIYGLNSLYLNTYIISNNLNKNYNNNHIFNYKY
NNNNFNKMSYNNKYKLPNFTGMNINNSICLNNIYKLYKHLIKDKFRYFYG
LYSELTQPIKGTGKOGOREGEMWALEAGASYLFPKEFTYKSDDIKSRKILKNY
LFNFKIKNTFISETFLLKELQSLAINIEAFICFNDNNLENLPINIIY"
10108..11835
/gene="rpoc"
10108..11835
/gene="rpoc"
/codon_start=1
/protein_id="CAA64573.1"
/db_xref="GI:1171588"
/db_xref="SWISS-PROT:P21422"
/translation="MIHNNINFIGLKLNILNPKQIIKWSSIFYKKNKIIIGEVILPNT
INFTGLPILNGLFCEKIFDYMYKNCNCKNNMYNNNFSPFLYCKFCCKNKLINNR
KYKLGFTPLNIPILHWLTGPKVASLLNKNVYFLKFLIYKYFTSNIKYQYVY
NKLPSKINLYKKYKNIQVLFSHNIIYKKLONILLTELLNKKELLINNKYNNKY
LYKKNLENFLINNIKPNWIFDLPLPAGLRPYPIYNNSTYIISTINENYRLIIL
NNKLYWLYRNIIFFEFIERKLIQDLIDYLLNKLILKNNTFNFSKTFQGGY
STIKYKLLGRVDFSGSVITNPSTIYNNIGLPIYISINLKPDLINILKYSKLNKI
IFKSLINKNLIIFQKFNRLNQFIINRAPTLHRNMPKPLLETEGYSLAFYPL
GCTSFNADFDGDOMSIFPLIKTFSFESINLNFKNIIISPSNNKLFNSLOYVKGI
NTLILANNVNLNFIYENSIEKIVYEYNNNLIIFLNLVWIKYINNNTIYILTINRI
ILNLYMYIY"
11844..14725
/gene="rpob"
join(11844..13418,13418..14725)
/gene="rpob"
/note="frameshift"
/codon_start=1
/protein_id="CAA64574.1"
/db_xref="GI:1171589"
/db_xref="SPTREMBL:Q25802"
/translation="MYFVFNKNIKLEKKLLIIFKYNISFKILHELLYLGYESFL
YNSLNKDFSNFIYLLIYKKNINNNKYVEIKNNYINVLNNYVYLYKVNKIQG
ILNNNLYKNIPYNSLFLFNKKIKIKYSOLOOLIGYKGISNKGIMKIEKPVINNY
ILNLIYIILSCYSKKGIIIDTALKTADSGYGLKRLINITSFIKELCKSPFLLK
YILNMDIYGNILILRFLQNNLNINNGFTYTKNTYITKTYLKLKLLNLYNR
NIYLNKISVYLCNITNNTCLNLYKQYKYNLQGHIGVISSEISSEPSQMVLRTFH
```

ASSILKDKFNFKYLYKLYKLNKLNKFKLIINFKKYVINKFNLIFLNMKILNLYN
NHLFEKYILOQYIKCNFYINSISKNFYNNNNIIKYNVINYNYNQLLIKIN
IHNKLYEITVLYYHYHFKYNNKGIILNNNNNKYNVIFYLINFYNFPSNYK
IYNNNFNSNYPKKMFILKNFNIOILNKLKVNNIYFYVKEKLFYLNIN
NIIKKYLFYKTYKLEFKKYNFLYELFKYNNKYLLNKNLYLIYNNYI
KYLYKNINILFKNLFYNNNFHIIHIIKNNYIYNNNNLYKNNKLNINNL
LYNKLFPYNNNNYLNIDITIGLOSINIFENKNIKONIFFISNNIYIFYKY
YLNLIYINIKYNNIYHKLQSYIFEDISSLYSGSLHTEFYFNKLY
YFRLKSYNIQATKSYIYVNIILIESILKOYSQNIYLPISYFELIKMLSCIK
IISNFKIFYNDIISLOLINIYSLNKHVIYKYPEIILGITKSLILANGSLFNI
SQNTFKIISLILNKNKIDWLDIKSKILITDLPVNGWYRYLVN"
14738...15421
/gene="rps2"
14738...15421
/gene="rps2"
/codon_start=1
/protein_id="CA64575.1"
/db_xref="GI:1171590"
/db_xref="SPTREMBL:O25803"
/translation="MFITFDNLKSKIYIGNIYKNIYFDNYKYIKRNFYCNILNFTF

Query Match 91.4%; Score 543; DB 3; Length 15421;
Best Local Similarity 95.8%; Pred. No. 3.3e-66;
Matches 569; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

Qy 1 GTATCGCTTTAATAGGCGAAGACAGACTTACCCCTTAAACATCTACTGCTTAGGATCGGA 60
Db 4704 GTATCGCTTTAATAGGCGAAGACAGACTTACCCCTTAAACATCTACTGCTTAGGATCGGA 4645
Qy 61 TAAGCCGACATCGAGGTCGCAACCTTTTCGTCGAATGAGCTCTCGGAAAGATTAGCC 120
Db 4644 TAAGCCGACATCGAGGTCGCAACCTTTTCGTCGAATGAGCTCTCGGAAAGATTAGCC 4585
Qy 121 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTAATCG 180
Db 4584 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTAATCG 4525
Qy 181 GATCATTAAAGACCGACATTATCTCTGTTTAAATTTGTAATTTTACAGTTAATTAATAT 240
Db 4524 GATCATTAAAGACCGACATTATCTCTGTTTAAATTTGTAATTTTACAGTTAATTAATAT 4465
Qy 241 T-TATCTTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 299
Db 4464 TATATCTTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4405
Qy 300 CCCAGTCAAACTATCTATAATAATGTTTAAATAATTTGTTTAAATAATTTTATAAGAA 359
Db 4404 CCCAGTCAAACTATCTATAATAATGTTTAAATAATTTGTTTAAATAATTTTATAAGAA 4345
Qy 360 TTTATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 419
Db 4344 TTTATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4285
Qy 420 TACTACTTCCCATTTATCTGTTATATATATATATATATATATATATATATATATATATAA 479
Db 4284 TATGTTTCCCATTTATCTGTTAA 4225
Qy 480 GCTTCATAGGCTCTTCTGCTCAATATAAGAAATCTGCAATCTTCACAGATAATTTTATT 539
Db 4224 GCTTCATAGGCTCTTCTGCTCAATATAAGAAATCTGCAATCTTCACAGATAATTTTATT 4165
Qy 540 TCATTAAAGATTTTTTTTAAAGACGACATTTTAAAGTCGTTTACATCTTTTCATCGAGGTC 594
Db 4164 TCATTAAAGATTTTTTTTAAAGACGACATTTTAAAGTCGTTTACATCTTTTCATCGAGGTC 4110

RESULT 10
PFU97560/c
LOCUS
DEFINITION
Accession
Version

malaria parasite P. falciparum.
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 595)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct PCR amplification and sequence analysis of extrachromosomal
Plasmodium DNA from dried blood spots
JOURNAL
Acta Trop. 68 (1), 105-114 (1997)
MEDLINE
98013247
PubMed
9352006
REFERENCE
2 (bases 1 to 595)
AUTHORS
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
TITLE
Direct Submission
JOURNAL
Submitted (15-APR-1997) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore
FEATURES
source
1..595
/organism="Plasmodium falciparum"
/isolate="primary isolate pf18/s"
/db_xref="taxon:5833"
/note="extrachromosomal plasmid"
<1..>595
/product="large subunit ribosomal RNA"
BASE COUNT 231 a 71 c 91 g 202 t
ORIGIN
Query Match 91.1%; Score 541.4; DB 3; Length 595;
Best Local Similarity 95.5%; Pred. No. 1.1e-65;
Matches 568; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
Qy 1 GTATCGCTTTAATAGGCGAAGACAGACTTACCCCTTAAACATCTACTGCTTAGGATCGGA 60
Db 595 GTATCGCTTTAATAGGCGAAGACAGACTTACCCCTTAAACATCTACTGCTTAGGATCGGA 536
Qy 61 TAAGCCGACATCGAGGTCGCAACCTTTTCGTCGAATGAGCTCTCGGAAAGATTAGCC 120
Db 535 TAAGCCGACATCGAGGTCGCAACCTTTTCGTCGAATGAGCTCTCGGAAAGATTAGCC 476
Qy 121 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTAATCG 180
Db 475 TGTATCCCTAGAGTAACCTTTTATCGTTAAGCGATAATTTTATTAATAATTAATCG 416
Qy 181 GATCATTAAAGACCGACATTATCTCTGTTTAAATTTGTAATTTTACAGTTAATTAATAT 240
Db 415 GATCATTAAAGACCGACATTATCTCTGTTTAAATTTGTAATTTTACAGTTAATTAATAT 356
Qy 241 T-TATCTTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 299
Db 355 TATATCTTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 296
Qy 300 CCCAGTCAAACTATCTATAATAATGTTTAAATAATTTGTTTAAATAATTTTATAAGAA 359
Db 295 CCCAGTCAAACTATCTATAATAATGTTTAAATAATTTGTTTAAATAATTTTATAAGAA 236
Qy 360 TTTATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 419
Db 235 TTTATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 176
Qy 420 TACTACTTCCCATTTATCTGTTATATATATATATATATATATATATATATATATATATAA 479
Db 175 TATGTTTCCCATTTATCTGTTATATATATATATATATATATATATATATATATATATAA 116
Qy 480 GCTTCATAGGCTCTTCTGCTCAATATAAGAAATCTGCAATCTTCACAGATAATTTTATT 539
Db 115 GCTTCATAGGCTCTTCTGCTCAATATAAGAAATCTGCAATCTTCACAGATAATTTTATT 56
Qy 540 TCATTAAAGATTTTTTTTAAAGACGACATTTTAAAGTCGTTTACATCTTTTCATCGAGGTC 594
Db 55 TCATTAAAGATTTTTTTTAAAGACGACATTTTAAAGTCGTTTACATCTTTTCATCGAGGTC 1

```

RESULT 11
PFU97558/c
LOCUS
DEFINITION
Plasmodium falciparum extrachromosomal plasmid large subunit
ribosomal RNA gene, partial sequence.
ACCESSION
U97558
VERSION
U97558.1 GI:4100396
KEYWORDS
malaria parasite P. falciparum.
SOURCE
Plasmodium falciparum
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 595)
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
Direct PCR amplification and sequence analysis of extrachromosomal
Plasmodium DNA from dried blood spots
Acta Trop. 68 (1), 105-114 (1997)
MEDLINE
98013247
PUBMED
9352006
REFERENCE
2 (bases 1 to 595)
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
Direct Submission
Submitted (15-APR-1997) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore
FEATURES
Location/Qualifiers
source
1..595
/organism="Plasmodium falciparum"
/isolate="primary isolate pf19/i"
/db_xref="taxon:5833"
/note="extrachromosomal plasmid"
<1..>595
/product="large subunit ribosomal RNA"
BASE COUNT 233 a 72 c 89 g 201 t
ORIGIN
Query Match 90.9%; Score 539.8; DB 3; Length 595;
Best Local Similarity 95.3%; Pred. No. 1.8e-65;
Matches 567; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
* QY 1 GATCGCTTTAATAGCGACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 60
|||||
Db 595 GATCGCTTTAATAGCGACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 536
QY 61 TAAGCGACATCGAGGTGCGCAACCTTTTCGTCATATGACACTCTCGGAAAGATTAGCC 120
Db 535 TAAGCGACATCGAGGTGCGCAACCTTTTCGTCATATGACACTCTCGGAAAGATTAGCC 476
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
Db 475 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 416
QY 181 GATCATTAAGACCGACATTAATCTCTGTTTAATTTGTAATTTTACAGTTAATTATATAT 240
Db 415 GATCATTAAGACCGACATTAATCTCTGTTTAATTTGTAATTTTACAGTTAATTATATAT 356
QY 241 T-TATCTTTATATAATAATAATAACATGTACACCTCCGTTTATATAGGAGGAGACCG 299
Db 355 TATATCTTTATATAATAATAATAACATGTACCTCCTCTGTTTATATATAGGAGGAGACCG 296
QY 300 CCCAGTCAAACTATCTTATAATATGTTTAAATTTTGTATAAAATTTTATATAAGAA 359
Db 295 CCCAGTCAAACTATCTTATAATATGTTTAAATTTTGTATAAAATTTTCTATAAGAA 236
QY 360 TTTATATATATAATAATGGTATTTTCATTAACAATTTACATTTATTTCCAAAAAATAATAT 419
Db 235 TTTATATGTAATAATAAATGGTATTTTCATTTTAACTAAATTTTCCAAAGAAATAATAT 176
QY 420 TACTACTTCCCATTTATCTATGTTATATATATATATATATTTTCAATATCTATTAAGTAA 479
Db 175 TATTGTTTCCCATTTATCTATGTTAAATATATATATATTTTTCATTTATTTATTAAGTAA 116
QY 480 GCTTCATAGGGCTTCTGCTCTCTAATATAAAGAACTCGCATCTTCACGATAATTTTATT 539
|||||

```

```

Db 115 GCTTCATAGGCTCTTCTGCTCTAATAAGAAATCTGTATCTTCACAGATAATTTTATT 56
QY 540 TCATTAAAGATTTTTTTTAAAGACAGCATTTAAGTCGTTACATCTTTCATCGAGGTC 594
|||||
Db 55 TCATTAAAGATTTTTTTTAAAGACAGCATTTAAGTCGTTACATCTTTCATCGAGGTC 1
|||||
RESULT 12
PFU97559/c
LOCUS
DEFINITION
Plasmodium falciparum extrachromosomal plasmid large subunit
ribosomal RNA gene, partial sequence.
ACCESSION
U97559
VERSION
U97559.1 GI:4100397
KEYWORDS
malaria parasite P. falciparum.
SOURCE
Plasmodium falciparum
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 594)
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
Direct PCR amplification and sequence analysis of extrachromosomal
Plasmodium DNA from dried blood spots
Acta Trop. 68 (1), 105-114 (1997)
MEDLINE
98013247
PUBMED
9352006
REFERENCE
2 (bases 1 to 594)
Tan,T.M.C., Nelson,J.S., Ng,H.C., Ting,R.C.Y. and Kara,U.A.K.
Direct Submission
Submitted (15-APR-1997) Institute of Molecular and Cell Biology,
National University of Singapore, 10 Kent Ridge Crescent S119260,
Singapore
FEATURES
Location/Qualifiers
source
1..594
/organism="Plasmodium falciparum"
/isolate="primary isolate pf20/l"
/db_xref="taxon:5833"
/note="extrachromosomal plasmid"
<1..>594
/product="large subunit ribosomal RNA"
BASE COUNT 231 a 71 c 90 g 202 t
ORIGIN
Query Match 89.7%; Score 532.6; DB 3; Length 594;
Best Local Similarity 95.6%; Pred. No. 1.8e-64;
Matches 569; Conservative 0; Mismatches 24; Indels 2; Gaps 2;
QY 1 GATCGCTTTAATAGCGACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 60
Db 594 GATCGCTTTAATAGCGACAGACTTACCCCTTAAACATACACTACTGCTTAGGATCGGA 535
QY 61 TAAGCGACATCGAGGTGCGCAACCTTTTCGTCATATGACACTCTCGGAAAGATTAGCC 120
Db 534 TAAGCGACATCGAGGTGCGCAACCTTTTCGTCATATGACACTCTCGGAAAGATTAGCC 475
QY 121 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 180
Db 474 TGTATCCCTAGAGTAACCTTTTATCCGTTAAGCGATAATTTTATTATTAATAATTATCG 415
QY 181 GATCATTAAGACCGACGATTAATCTCTGTTTAATTTGTAATTTTACAGTTAATTATATAT 240
Db 414 GATCATTAAGACCGACGATTAATCTCTGTTTAATTTGTAATTTTACAGTTAATTATATAT 355
QY 241 T-TATCTTTATATAATAATAATAACATGTACACCTCCGTTTATATAGGAGGAGACCG 299
Db 354 TATATCTTTATATAATAATAATAACATGTACTCCTCCGTTTATATAGGAGGAGACCG 295
QY 300 CCCAGTCAAACTATCTTATAATATGTTTAAATTTTGTATAAAATTTTATATAAGAA 359
Db 294 CCCAGTCAAACTATCTTATAATATGTTTAAATTTTGTATAAAATTTTCTATAAGAA 235
QY 360 TTTATATATATAATAATGGTATTTTCATTAACAATTTACATTTATTTCCAAAAAATAATAT 419
Db 234 TTTATATATAATAATAATGGTATTTTCATTTTAACTAAATTTTCCAAAGAAATAATAT 176

```


Db	416	GGATCATTAAGACCGACATTTATCTGTTTAATTTGTAATTTTACAGTTAATTATATA	357
Qy	240	TT-TATCTTTTATATAATAACATTTGTACACCTCGGTTTTTATATAGGAGGAGACC	298
Db	356	TTATATCTTTATATAATAATAACATTTGTACTCTCGGTTTATATATAGGAGGAGACC	297
Qy	299	GCCCCAGTCAAACTATCTATATAAATTTGTTAAAAAATTTGTTATAAAAATTTTATAAGA	358
Db	296	GCCCCAGTCAAACTATCTATATAAATTTGTTAAAAAATTTGTTATAAAAATTTTCTATAAGA	237
Qy	359	ATTTATATATATAATAAATGGTATTTTCATTAACAATTTACATATTTCCAAAATAATAATA	418
Db	236	ATTTATATATAATAAATGGTATTTTCATTTTAACTAAATTTATTTCCAGAAAATAATAATA	177
Qy	419	TTACTACTTCCCATTTATTTCTATGTATATATATATATTTTCAATATCTATTATAAGTAA	478
Db	176	TTAATGTTTCCCATTTATTTCTATGTATATATATATATTTTCATTTATTTAATAAGTAA	117
Qy	479	AGCTTCATAGGGTCTTTCTGTCCTTAATATAAGAAAATCGCATCTTCACAGATAAATTTTAT	538
Db	116	AGCTTCATAGGGTCTTTCTGTCCTTAATATAAGAAAATCTGTCTTCACAGATAAATTTTAT	57
Qy	539	TTCTATTAAGATTTTTTTTAAAGACAGCATTTAAGTCGTTACATCTTTCATGCGAGTC	594
Db	56	TTCTATTAAGATTTTTTTTAAAGACAGCATTTAAGTCGTTACATCTTTCATGCGAGTC	1
RESULT 15			
PFU97557/c			
LOCUS			
DEFINITION	PFU97557	595 bp	DNA linear INV 15-MAR-2001
ACCESSION	Plasmodium falciparum	extrachromosomal	plastid large subunit
VERSION	ribosomal RNA gene, partial sequence.		
KEYWORDS	U97557		
SOURCE	U97557.1	GI:4100395	
ORGANISM	malaria parasite P. falciparum.		
REFERENCE	Plasmodium falciparum		
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
TITLE	1 (bases 1 to 595)		
JOURNAL	Tan, T.M.C., Nelson, J.S., Ng, H.C., Ting, R.C.Y. and Kara, U.A.K.		
MEDLINE	Direct PCR amplification and sequence analysis of extrachromosomal		
PUBMED	Plasmodium DNA from dried blood spots		
REFERENCE	Acta Trop. 68 (1), 105-114 (1997)		
AUTHORS	98013247		
TITLE	2 (bases 1 to 595)		
JOURNAL	Tan, T.M.C., Nelson, J.S., Ng, H.C., Ting, R.C.Y. and Kara, U.A.K.		
REFERENCE	Direct Submission		
TITLE	Submitted (15-APR-1997) Institute of Molecular and Cell Biology,		
JOURNAL	National University of Singapore, 10 Kent Ridge Crescent S119260,		
FEATURES	Singapore		
SOURCE	Location/Qualifiers		
	1..595		
	/organism="Plasmodium falciparum"		
	/isolate="primary isolate pf11/p"		
	/db_xref="taxon:5833"		
	/note="extrachromosomal plastid"		
	<1..>595		
	/product="large subunit ribosomal RNA"		
BASE COUNT	229 a	77 c	93 g 196 t
ORIGIN			
Query Match	89.0%;	Score 528.6;	DB 3; Length 595;
Best Local Similarity	94.1%;	Pred. No. 6.2e-64;	
Matches 560;	Conservative	0; Mismatches 34;	Indels 1; Gaps 1;
Qy	1	GTATCGCTTTAATAGGCGAAGACAGACTTACCTTTAAAACTACTACTGCTTAGATCGGA	60
Db	595	GTATCGCTTTAATAGGCGAAGACAGACTTACCTTTGGGCGATCTACTGCTTTAGATCGGA	536
Qy	61	TAGCCGACATCGAGGTGCGAAACCTTTTCGTCAATGATGACTCTCGAAAGATTAGCC	120

